Appendix 2 Tables of abstracted documents

Table A2.1 Workforce to patient ratios

ID, origin, authors (year)	7, USA, Aiken, L.H., Clarke, S.P., Sloane, D.M. *et al.* (2002)		
Aims	To determine the association between the nurse:patient ratio and patient mortality, failure to rescue among surgical patients and factors related to nurse retention. *Workforce:* Registered Nurses, secondary care *Feature:* Nurse:patient ratio (calculated as the mean patient load across all staff RNs who reported having responsibility for at least 1 but <20 patients on the last shift they worked, regardless of their specialty) *Outcome:* 30-day surgical mortality and failure to rescue (deaths of surgical patients within 30 days following complications, identified by scanning discharge abstracts for ICD-9-CM codes). Nurse reported job dissatisfaction and job-related burnout were studied but not in relation to patient outcomes.		
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, cross-sectional 2 Adult general hospitals with patients undergoing general surgical, orthopaedic or vascular procedures who were aged between 20 and 85 3 232,342 patients 4 30 days after surgery, in hospital 5 Discharge abstracts obtained from the Pennsylvania Health Care Cost Containment Council were merged with Pennsylvania vital statistics records to identify patients who died within 30 days of hospital admission to control for timing of discharge as a possible source of variation in hospital outcomes. Hospital characteristics were derived from the 1999 AHA Annual Survey and 1999 Pennsylvania DoH Hospital survey. Data collected from April 1 1998 to November 30 1999.		
Results Quantitative results		**Mortality**	**Failure to rescue**
	Unadjusted	OR = 1.14 (1.08–1.19) *p*-value <0.001	OR = 1.11 (1.06–1.17) *p*-value = 0.004
	Adjusted for patient characteristics	OR = 1.09 (1.04–1.13) *p*-value <0.001	OR = 1.09 (1.04–1.13) *p*-value = 0.001
	Adjusted for patient and hospital characteristics	OR = 1.07 (1.03–1.12) *p*-value <0.001	OR = 1.07 (1.02–1.11) *p*-value <0.001
	OR = Odds ratios indicate the risk associated with an increase of 1 patient per nurse. Direct standardisation techniques were used to predict excess deaths in all patients and in patients with complications that would be expected if the patient:nurse ratio were at various levels in the California staffing mandate debates. Additional deaths or failures associated with a ratio of 10 patients per nurse were not calculated because of the limited number of hospitals in the sample staffed at that level. **Ratio of patients per nurse** 6:4 = 2.3 (1.1–3.5) additional deaths per 1000 patients and 8.7 (3.9–13.5) additional deaths per 1000 patients with complications. 8:6 = 2.6 (1.2–4.0) additional deaths per 1000 patients and 9.5 (3.8–15.2) additional deaths per 1000 patients with complications. 8:4 = 5 (2.4–7.6) additional deaths per 1000 patients and 18.2 (7.7–28.7) additional deaths per 1000 patients with complications. When taken together, the impacts of staffing on patient and nurse outcomes suggest that by investing in RN staffing hospitals may reduce both preventable mortality and turnover rates in hospitals by reducing burnout and job dissatisfaction, major precursors of job resignation. The overall result remains unchanged when adjustments were made for patient:licensed nurse ratios and patient:unlicensed assistive personnel ratios.		

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Controlled for demographic characteristics of patients, nature of hospital admission, co-morbidities and relevant interaction terms using 133 variables, including age, sex, surgery types, and dummy variables indicating the presence of chronic pre-existing health conditions and interaction terms. 2 Adjusted for hospital size, teaching status and technology. *Size:* Small <100 beds, Medium 101–250 beds and Large >251 beds. *Teaching status:* Ratio of resident physicians and fellows to hospital beds. Hospitals with no postgraduate trainees (non-teaching) were contrasted with those that had a 1:4 or smaller trainee:bed ratios. *Technology:* High: Hospitals with facilities for open heart surgery and/or major transplants. Contrasted with other hospitals. 3 Yes 4 52% response rate from the survey; 168 of the 210 hospitals had discharge data for surgical patients in the targeted DRGs as well as AHA data and survey data from 10 or more staff nurses. 6 of the excluded hospitals were Veterans Affairs Hospitals, which do not report discharge data to the state. 26 hospitals were excluded because variables were missing from the administrative or patient outcomes and therefore could not be matched to the surveys. 10 small hospitals (most <50 beds) were excluded as fewer than 10 nurses responded to the survey. 5 50% random sample of RNs were sent a survey who were on the Pennsylvania Board of Nursing rolls and resided in the state. 6 Pennsylvania
Commentary	The failure to weight nurses' responses by the type of unit or shift on which they work could introduce measurement errors in the average patient:nurse ratio used as the independent variable. A hospital with a larger proportion of nurses responding who work in intensive care settings will have a higher patient:nurse ratio than a hospital where more medical–surgical unit nurses respond even if the unit-by-unit staffing ratios are identical. The response rate of 52% may compare favourably with rates seen in other voluntary surveys, and the sample closely resemble those participating in the National Sample Survey of RNs, but the sample may be less representative of nurses working in hospitals. The number of responses from some hospitals is low. There were 50 responses from fewer than half the hospitals and this could introduce response bias.
Research implications	Longitudinal data sets are needed to exclude the possibility that low hospital nurse staffing is the consequence, rather than the cause, of poor patient and nurse outcomes. Is turnover and retention of staff linked with mortality and failure to rescue? How many nurses are needed to care for patients? Is there a maximum ratio of patients per nurse above which hospitals should not exceed? What do registered nurses do and when, how, and where do they add value to the quality and outcomes for patients?

ID, origin, authors (year)	229, USA, Amaravadi, R.K., Dimick, J.B., Pronovost, P.J. and Lipsett, P.A. (2000)
Aims	To determine if having a night-time nurse:patient ratio (NNPR) of 1 nurse caring for 1 or 2 patients versus 1 nurse caring for 3 or more patients in the ICU is associated with clinical and economic outcomes following oesophageal resection. *Workforce:* Nurse, ICU *Feature:* Night-time nurse:patient ratio (<1:2 or >1:2) *Outcome:* In-hospital mortality, hospital LOS and complications after esophagectomy
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observational including survey 2 All adults discharged from Maryland hospitals with a primary procedure code for oesophageal resection were included. 3 353 patients, 32 acute-care hospitals 4 In hospital 5 Patient data were obtained from hospital discharge data available from the Maryland Health Service Cost Review Commission (HSCRC). Data were collected on nurse staffing in 1996 from a survey on ICU organisational characteristics, which were mailed to physician ICU directors; the survey included 32 items on organisational characteristics, ICU physician and nurse characteristics and processes of care. Data collection: 1994–1998
Results Quantitative results	**Complication** / **NNPR >1:2 (%)** / **NNPR <1:2 (%)** / **Odds ratio** / ***p*-value** Pneumonia: 8 / 16 / 2.4 (1.2–4.7) / 0.012 Reintubation: 12 / 25 / 2.5 (1.4–4.5) / 0.001 Aspiration: 22 / 25 / 1.2 (0.7–2.0) / 0.5 Septicaemia: 1.8 / 6.2 / 3.7 (1.1– 12.5) / 0.04 Postoperative infection: 4 / 5.5 / 1.4 (0.5–3.8) / 0.5 Myocardial infarction: 0.9 / 0.8 / 0.9 (0.08–9.7) / 0.9 Cardiac arrest: 0 / 0.8 / 1.2 (0.6–2.2) / 0.6 Surgical complications: 8 / 17 / 1.9 (0.9–3.8) / 0.08 Acute renal failure: 2.7 / 5.5 / 2.1 (0.7–6.4) / 0.2 *In-hospital mortality:* The unadjusted mortality rate for patients with a NNPR <1:2 versus those with a NNPR >1:2 was 15% vs. 5.6% (p = 0.009). There was no significant difference in the risk of in-hospital mortality between patients with the 2 staffing ratios (OR = 0.7, 0.3 – 2.0) after adjustments. *In-hospital LOS*: The median LOS for patients with a NNPR <1:2 vs. NNPR >1:2 was 15 days vs. 9 days (IQR = 1.8–13, p <0.001). There was a 39% increase in LOS for patients with an NNPR <1:2 compared with an NNPR >1:2 (CI = 19–61%; p <0.001). Using multi-level hierarchical modeling (clustering) the point estimate for NNPR <1:2 remained the same but the confidence interval expanded to include zero (–8–109%; p = 0.11)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 *Patient characteristics:* Age, sex and race (white or non-white); co-morbidity: yes or no for up to 12 secondary discharge diagnoses and 14 secondary procedures diseases adjusted using the Romano-Charlson co-morbidity index; type of operation: (transhiatal, transthoracic, unspecified) and operating physician; nature of admission: (elective, urgent or emergent) 2 *Organisational characteristics:* volume (low = <25 cases per year for hospitals and <10 for surgeons or high) and vital status at discharge 3 Yes 4 Unit data were unavailable for three centres 5 No 6 Maryland

Commentary	The sample size was insufficient and therefore when hierarchical modelling was performed the CI widened. Errors in the coding of co-morbid diseases and complications could have affected the results. The different methods of obtaining patient and ICU data could have been a source of bias. In order to minimise this bias, the data abstraction was blinded to hospital name, ICU characteristics and patient outcome. A prospective scoring system was not used such as the APACHE score. Pre-ICU and post-ICU care could not be adjusted for. The validity of the survey instrument could be a limitation as the data was collected in 1996 and the ratios may have changed before or after the survey. There were no questions regarding nursing experience, ICU nursing experience or nursing certification.
Research implications	Nursing care takes on an increased importance at night, when physician and ancillary service staffing is typically decreased. As the number of patients each nurse cares for increases, the time that can be devoted to each patient decreases.

ID, origin, authors (year)	87, USA, Dang, D., Johantgen, M.E., Pronovost, P.J. *et al.* (2002)
Aims	To examine the association between intensive care unit nurse staffing and the likelihood of complications for patients undergoing abdominal aortic surgery *Workforce:* Nurse, ICU *Feature:* Nurse:patient ratios (low = 1:3 or > on the day and night shift, medium = 1:3 or > on either the day or night shift, but not both and high intensity = 1:2 or < on the day and night shifts) and nurse:patient ratios during the day (more = 1:1 or 1:2 and fewer = 1:3 or 1:4) *Outcome:* Medical complications of patients undergoing abdominal aortic surgery: cardiac, respiratory and other
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, retrospective chart review and secondary analyses of survey data 2 Population drawn from inpatient hospital stays for all patients undergoing abdominal aortic surgery in Maryland, patients aged 30 or over who were discharged from a Maryland hospital who had a principal procedure code for abdominal aortic surgery 3 2606 patients, 38 acute-care hospitals 4 In-hospital 5 Maryland Health Discharge Data Set maintained by the Maryland Health Services Cost Review Commission (HSCRC). Data were collected on nurse staffing from a survey on ICU organisational characteristics, which was mailed to physician ICU directors; the survey included 32 items on organisational characteristics and asked: what is the average nurse:patient ratio in the ICU during the daytime and then the night-time? Data collection: January 1994 and December 1996.
Results Quantitative results	**Overall ratios** Hospital and surgeon volume, hospital and ICU beds, and critical paths and protocols were all highly correlated with each other. Variables in each pair that had the largest variance were excluded from the multivariate analysis.

Staffing Intensity	**n**	**Cardiac (n=341) Odds ratio (95% CI)**	**Respiratory (n=787) Odds ratio (95% CI)**	**Other (n=221) Odds ratio (95% CI)**
High (referent category)	1600	1.00	1.00	1.00
Medium	586	1.78 (1.16-272)	1.03 (0.78–1.38)	1.74 (1.15–2.63)
Low	420	1.34 (0.82–2.17)	2.33 (1.50–3.60)	1.13 (0.73–1.75)
Hosmer-Lemeshow Statistic		X^2=9.70 p=0.29	X^2=12.15 p=0.14	X^2=7.40 p=0.49

	Cardiac complications after procedure	**Pulmonary insufficiency after surgery**	**Mechanically ventilated after 96 hours**	**Re-intubated**
High (referent)	1.00	1.00	1.00	1.00
Medium	2.10 (1.26–3.50)	–	–	–
Low	–	5.11 (2.89–9.04)	2.39 (1.55–3.69)	2.09 (1.47–3.03)

Patients cared for with medium-intensity staffing were more likely to have cardiac complications and other complications than patients cared for on units with high-intensity staffing. Patients cared for on units with low-intensity staffing were more than twice as likely to have respiratory complications as patients on units with high-intensity staffing. Volume of cases was significantly related to respiratory complications, patients cared for in low-volume hospitals were 82% more likely to have a respiratory complication (OR = 1.82, 1.25–2.67) compared to high-volume hospitals.

Ratios during the day

Fewer versus more ICU nurses per patient were independently associated with an increased risk of medical complications.

Complications	Hospitals with fewer ICU nurses (n=478) %	Hospitals with more ICU nurses (n=2128) %	Relative risk	
			Crude	Adjusted
Any complications	47	34	1.4 (1.2–1.5)	1.7 (1.3–2.4)
Any medical complications	43	28	1.5 (1.4–1.7)	2.1 (1.5–2.9)
Pulmonary insufficiency after procedure	24	9	2.6 (2.1–3.2)	4.5 (2.9–6.9)
Re-intubation	21	13	1.5 (1.3–1.8)	1.6 (1.1–2.5)
Cardiac complications after procedure	15	10	1.4 (1.1–1.7)	1.3 (0.8–1.8)
Acute renal failure	6	4	1.3 (0.8–1.9)	1.6 (0.9–2.7)
Septicemia	4	3	1.4 (0.8–2.1)	1.9 (1.9–3.9)
Acute myocardial infarction	4	3	1.5 (0.8–2.4)	1.5 (0.9–2.2)
Cardiac arrest	2	1	1.4 (0.6–3.0)	1.7 (0.7–1.5)
Any surgical complication	10	11	0.9 (0.6–1.4)	0.7 (0.4–1.5)
Surgical complications after procedure	8	9	0.9 (0.6–1.2)	1.0 (0.6–1.4)
Surgical E codes	1	0	2.2 (0.4–10.5)	Insufficient data
Re-operation for bleeding	2	3	0.8 (0.4–1.6)	1.2 (0.4–3.5)

Quality appraisal

1 Case mix adjustment
2 Other adjustment
3 Uniform data collection
4 Participant follow-up
5 Random sampling
6 Geographical dispersal

1 *Patient characteristics:* Age, sex and race (white or non-white); co-morbidity: yes or no for 10 diseases in the Romano-Charlson co-morbidity index; severity: type (ruptured or non-ruptured aneurysm) and nature of admission (elective, urgent or emergent).
2 *Organisational characteristics:* Number of hospital and ICU beds, volume (low = <36 cases per year for hospitals and <8 for surgeons or high) of aortic surgery performed during study period, type of unit, full-time medical director and nurse manager, RN attendance at daily rounds, and use of written protocols and critical paths for abdominal aortic surgery patients.
3 Yes
4 9 patients under 30 were excluded because they had suffered an injury to a blood vessel; 7 hospitals did not respond to the survey and were excluded (no differences in hospital and patient characteristics in responders and non-responders)
5 No
6 Maryland

Commentary	The increased likelihood of all complications in ICUs may reflect a difference in the level of monitoring by nurses or possibly an insufficient number of nurses to perform interventions such as pulmonary hygiene, an aspect of care for which nurses are responsible. Although the nursing response rate for the survey was 83% the nurse manager may or may not have been involved in the survey completion, and responses may represent perceptions or experiences over time. In ICUs there is likely to be less variation in nurse staffing compared with general units. The complications selected were those that were likely to be influenced by nursing interventions and the level of nursing surveillance, recognising that nurses are not only team members. Unsure of the reliability of the coding of co-morbid diseases and complications. By using administrative databases there is an inability to determine whether the complication occurred in the ICU or general surgical unit. Optimal ICU organisation should address both physician and nurse staffing. The study is retrospective. Coding of the complications and co-morbid diseases in the HSCRC database may not be as accurate as the coding of the principal procedure. No systematic scoring system was used. No adjustments were made for differences in pre-ICU care, including surgical approach and type of anaesthesia, and post-ICU care. The ICU:patient ratio is a relatively crude measure of nursing surveillance; it is a complex variable that may be affected by staff mix, experience, training, certification, fatigue and nursing workload. Complications are influenced by a complex array of factors many of which may be unrelated to nursing. The study focused on only one surgical procedure in one state so the applicability of the findings to other procedures and other states is limited.
Research implications	Does this relationship between nurse staffing and complications occur through other process or contextual aspects of nursing units, such as the organisation of nursing services and the practice environment? Would the same conclusion hold if a more sensitive measure of nurse staffing, such as nursing hours per patient day obtained at the unit level, was used? Is it because there are not enough nurses to perform the procedures or not enough time to provide surveillance of the patients and catch early warning signs? What about skill mix, experience and staffing intensities of staff in the ICU? What is the optimal ICU nurse:patient ratio for an ICU with a given severity of illness? Do nurses who care for 3 or more patients in the ICU have less time than nurses who care for 1 or 2 patients to devote to patient care, especially preventive measures?

ID, origin, authors (year)	160, USA, Dimick, J.B., Swoboda, S.M., Pronovost, P.J. and Lipsett, P.A. (2001)
Aims	To determine if having 1 nurse caring for 1 or 2 patients (more nurses) versus 1 nurse caring for 3 or more patients (fewer nurses) in the ICU at night is associated with differences in clinical and economic outcomes after hepatectomy *Workforce:* Nurses, ICU *Feature:* Nurse–patient ratios at night (more nurses = 1:1 or 1:2, fewer nurses = 1:3 or more) *Outcome:* Mortality, LOS and complications after hepatectomy
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observational including survey 2 All adult (over 18) patients with a primary code for hepatic resection who were discharged from hospital during the study period in Maryland 3 556 adults and 33 acute-care hospitals 4 In-hospital 5 Maryland Health Discharge Data Set maintained by the Maryland Health Services Cost Review Commission (HSCRC). Data were collected on nurse staffing from a survey on ICU organisational characteristics, which was mailed to physician ICU directors. The survey listed 32 characteristics of ICU staffing by physicians and nurses and other aspects of ICU organisation and processes of care. Data collection: 1994–1998.
Results Quantitative results	The difference between the adjusted in-hospital mortality rate and LOS for the two groups was not significant. Fewer nurses were associated with an adjusted risk of in-hospital mortality of OR = 0.49 (0.18–1.29) and increase in hospital LOS of OR = 0.67 (–0.80–0.93). After adjustments were made the only complications that remained significantly associated with fewer ICU nurses at night was re-intubation. **Complication** / **More nurses** / **Fewer nurses** / **Odds ratio (95% CI)** / ***p-value*** Pneumonia / 2.8 / 4.2 / 1.4 (0.6–3.5) / 0.40 Reintubation / 1.9 / 10.8 / 5.7 (2.4–13.7) / <0.001 Pulmonary Failure / 1.6 / 5.8 / 3.6 (1.3–10.1) / 0.006 Aspiration / 12 / 7.5 / 0.62 (0.4–1.1) / 0.08 **Complication** / **More nurses** / **Fewer nurses** / ***p*-value** Septicemia / 2.7 / 5.4 / 0.27 Postoperative Infection / 2.9 / 3.0 / 0.96 Cardiac arrest / 0.6 / 0.8 / 0.90 Myocardial infarction / 6.6 / 1.2 / 0.27 Acute renal failure / 14.6 / 4.2 / 0.72
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 *Patient characteristics:* Age, sex and race (white or non-white); co-morbidity: yes or no for 14 diseases in the Romano-Charlson co-morbidity index; severity: type of operation and nature of admission (elective, urgent or emergent) 2 *Hospital characteristics:* Volume (low = <30 cases per year for hospitals and <10 for surgeons or high) of procedures performed during the study period by each hospital and each surgeon. 3 Yes 4 ICU survey data were unavailable for two of the centres performing hepatic resection 5 No 6 Maryland

Commentary	Unadjusted rates were reported in tables where significant but the adjusted rates were mentioned in the text. The authors report that as the number of patients each nurse cares for increases, the time available to devote to direct care of patients decreases thus allowing complications to go unnoticed. A potential bias in the study is the accuracy of coding in administrative databases. Measures were put in place to minimise the bias of data abstraction – blinding to hospital name, ICU characteristics and patient outcomes. The analysis might not account for factors that were not identified in the administrative database but were important at the patient or unit level. The validity of the survey instrument could be a limitation as the data were collected in 1996 and the ratios may have changed before or after the survey. There were no questions regarding nursing experience, ICU nursing experience or nursing certification. No investigation of the use of respiratory care practitioners or other ancillary staff. The study was retrospective and does not directly measure cause and effect and does not reflect the individual impact of a single nurse or physician.
Research implications	How do specific characteristics and processes of nursing care alter outcomes? Does the experience or training of the nurses on the night shift affect outcomes? Do the skill mix and grade mix of other staff alter the results of this study? Does the availability of the staff influence the outcomes?

ID, origin, authors (year)	278, USA, Fridkin, S.K., Pear, S.M., Williamson, T.H. *et al,* (1996)
Aims	To determine risk factors for central venous catheter-associated bloodstream infections during a protracted outbreak *Workforce:* Nurses, surgical intensive care unit (SICU) in a university-affiliated Veterans Affairs medical centre *Feature:* Nurse:patient ratios, nursing hours per month and SICU patient day *Outcome:* Central venous catheter-associated bloodstream infections (CVC-BSIs) or site infection rates and mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, case–control and cohort 2 Case patients: any patient hospitalised >=48 hours, in the SICU >=24 hours or who developed a laboratory-confirmed CVC-BSI during the outbreak period. A subset of these patients who received TPN and had their first laboratory confirmed CVC-BSI was chosen. Controls: selected randomly from a list of all SICU patients with a central venous catheter who did not develop a CVC-BSI and who had been in the SCIU >=24 hours during the outbreak period and remaining hospitalised on any ward >14 days. Cohort: compared SICU patients before and after the outbreak. 3 Case–control: 30 patients. The hospital had 230 beds. All SICU patients: 1760. 4 In-hospital 5 Case patients were identified by reviewing ICU, microbiology and infection control records and data were abstracted from patients' medical, respiratory therapy, pharmacy and nutrition records. Monthly SICU nursing hours worked and SICU patient census were obtained from nursing services to calculate nurse:patient ratios. January 1991 to September 1993.
Results Quantitative results	Nurse staffing changed significantly between the two periods. Although the number of SICU patient-days per month decreased, the number of hours worked by SICU nurses per month decreased proportionately more in the outbreak than the pre-outbreak period, resulting in an overall increase in the monthly average patient:nurse ratio. A high patient:nurse ratio was temporally associated with >=1 CVC-BSI occurring in the SICU. Furthermore, the number of SICU CVC-BSI correlated with the SICU patient:nurse ratio (Spearman's Rank Correlation Coefficient = 0.49 and $p<0.01$). An SICU patient:nurse ratio of at least 1.26 (the median value for the entire study period) was associated with the occurrence of >=1 CVC-BSI in the SICU (RR = 2.2, CI = 1.1–4.3). **Pre-outbreak period** / **Outbreak period** / ***p*-value** Nursing hours per month: 4297 / 3239 / <0.01 Nursing hours per SICU patient-day: 20.3 / 17.0 / <0.01 Patient:nurse ratio: 1.18 / 1.40 / <0.01 **Patient:nurse ratio** / **Nursing hours worked per patient-day** / **Adjusted* odds ratio** / **CI** 1 / 24 / 1 / – 1.2 / 20 / 3.95 / 1.07–14.54 1.5 / 16 / 15.6 / 1.15–211.4 2 / 12 / 61.5 / 1.23–3074 Looked at difference in groups for: age, gender, LOS (hospital and SICU), discharge diagnosis, outcome, characteristics before infection (i.e. assisted ventilation, intravenous therapy, TPN, central venous catheter use, operative procedures, concurrent illness), severity of illness as measured by the Acute Physiology and Chronic Health Evaluation (APACHE II) Score.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 *Adjusted for study period, total parenteral nutrition or assisted ventilation 2 No 3 Yes 4 Not stated 5 No 6 One hospital in Arizona

Commentary	Device-day data were lacking and therefore the results are not generalisable to other institutions. APACHE II scores could not be compared between all SICU patients in the outbreak and pre-outbreak periods. The criteria used to define catheter-site infections may have caused an underestimation of the rates of these types of infections, because only late stages of infection were counted. This may have introduced some bias. However, by focusing on laboratory-confirmed bloodstream infections and by showing a similar number of SICU patient blood cultures evaluated in the two periods the occurrence of selection bias was minimised. The experience of the SICU nurses was not taken into account. The case patients median score on the APACHE II scale was twice that of control patients.
Research implications	Further investigation of the relationship between infections and staffing levels adjusting for patient and hospital characteristics.

ID, origin, authors (year)	397, UK, Hunt, J. and Hagen, S. (1998)
Aims	To investigate the relationship between nurse staffing and patients *Workforce:* Registered Nurses, general hospital (teaching and non-teaching) *Feature:* Nurse:patient ratios (calculated by dividing the number of patients by the number of whole-time-equivalent (WTE) staff) *Outcome:* Mortality (death within 30 days after emergency admission to hospital) and re-admission rates (re-admitted within 30 days of discharge from hospital)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observational 2 Bank nurses and nurse teachers were excluded. 3 23 general hospitals and 67 staffed beds 4 30 days 5 Routine data for the nursing workforce were collected by the information services division (ISD) of the Common Services Agency for the NHS in Scotland on 30 September 1994. The number of occupied beds was taken as a proxy for the number of patients and was obtained from ISD for the year ending 31 March 1995. Patient outcomes were obtained from ISD for the period from April 1994 to March 1995.
Results Quantitative results	The NPR varied across the trusts, from 0.71 to 1.66 qualified nurses per patient, the mean being 1.21. Correlation coefficients were not found to be significant for either outcome. Using multiple regression re-admission rates were found to be significantly related to NPR but mortality was not. The regression coefficients indicated that increased re-admission rates were associated with lower total NPR. ANOVA for the association between mortality rate and NPR gave an F-value = 1.234 and *p*-value = 0.325* and for re-admission rates F-value = 3.608 and *p*-value = 0.032*.

	Unstandardised coefficients		**Standardised coefficients**		
	B	SE	Beta	T	*p*
Mortality	–0.314	0.548	–0.139	–0.573	–0.573
Readmission rates	–1.696	0.741	–0.484	–2.287	–0.034

Quality appraisal	
1 Case mix adjustment	1 *Outcomes:* standardised for age, sex and deprivation category
2 Other adjustment	2 *In multiple regression independent variable: (constant), trust teaching status, NPR, trust type (small/large)
3 Uniform data collection	3 Yes
4 Participant follow-up	4 Complete
5 Random sampling	5 No
6 Geographical dispersal	6 All general hospitals in Scotland
Commentary	Clinical grade was used as a proxy for skill mix as specialist qualifications are not routinely collected data. No adjustments were made for severity of illness. The authors state that the variation in the ratios appears to be due to the combined differences in ratios of grade D and E RNs to occupied staffed beds. Although no relationship was found between mortality and NPR a more detailed analysis of patient-level data, which considered simultaneously differences at speciality, trust and population levels, may have given different results. The authors state that using existing, routinely collected data was difficult. As the trusts were so different it was difficult to make comparisons between NPRs at trust level. The statistical task of separating out the influence of the individual input of confounding factors is difficult and data is not also available for all confounders. The relationship between teaching status and re-admission rates were different from other studies.
Research implications	Skill mix? Numbers of other nursing staff? Severity of illness?

ID, origin, authors (year)	1128, UK, Jarman, B., Gault, S., Alves, B. *et al.* (1999)
Aims	To ascertain hospital inpatient mortality in England and to determine which factors best explain variation in standardised hospital death ratios *Workforce:* GPs, nurses, mixed settings *Feature:* Ratios – ratio of hospital doctors to beds and GPs to heads of population *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental; retrospective 2 *Inclusion:* discharge records only (episodes that ended in discharge (alive or dead) from the hospital rather than transfer to the care of another consultant within the hospital) for primary diagnosis of one of 85 primary diagnoses which accounted for 80% of deaths. *Exclusion:* Community and speciality institutions, small hospitals (under 9000 admissions during the 4 years) and hospitals without accident and emergency units. 3 183 acute general hospital trusts 4 In-hospital 5 Three main sources: the NHS hospital episode statistics data system from 1991/2 to 1994/5, the national decennial census from 1991 and other routine NHS data such as hospital characteristics, hospital staffing levels and general practitioner distribution.
Results Quantitative results	Weighted multiple linear regression using two models: A included all admissions both elective and emergency; B included emergency admissions only. Stratified by age (using 10-year age groups), sex and the 85 primary diagnoses. Aggregate discharge data were taken from individual records and aggregated across each hospital. Community data were taken from geographical areas attributed from area of residence to each discharge (via postcode) and then averaged across discharges for each hospital. **Model A** After adjustments for percentage of emergency admissions, the best predictors of hospital mortality were numbers of hospital doctors per 100 hospital beds and general practitioners per 100,000 population. Higher hospital standardised mortality ratios were associated with lower numbers of hospital doctors per hospital bed and lower numbers of GPs per head of the population. A reduction in 5000 hospital deaths per year was associated with a 27% increase in hospital doctors or an 8.7% increase in general practitioners. **Variable** — **Regression coefficient (95% CI)** — ***p*-value** — **Mean** Number of hospital doctors per 100 hospital beds in 1994/5 — –0.47 (–0.64 to –0.30) — <0.001 — 25.4 (8.0) Number of general practitioners per 100,000 population — –0.67 (–1.05 to –0.30) — <0.001 — 54.6 (3.4) **Model B** At the 5% level of significance, the proportion of grade A nurses (auxiliary nurses in training) as a percentage of all hospital nurses and bed occupancy entered the model. High percentages of grade A nurses were associated with higher hospital standardised mortality ratios. **Variable** — **Regression coefficient (95% CI)** — ***p*-value** — **Mean** Number of hospital doctors per 100 hospital beds in 1994-5 — --0.51 (-0.65 to –0.38) — <0.001 — 25.4 (8.0)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 *Included in multiple regression analysis:* Aggregate discharge data: percentage of emergency cases; percentage of cases and deaths with co-morbidity of the 85 diagnoses leading to 80% of all deaths and combinations of those with the highest correlations with hospital standardised mortality ratios; percentage of cases and deaths with each of the top 15 diagnoses which account for 50% of deaths; percentage of cases with co-morbidity of the two or three conditions most highly correlated with hospital mortality; health authority where hospital located. Hospital data: hospital doctors per bed; percentage of nurses at grades A; bed occupancy; location (inner London and outer London); university teaching. Community attributed data: general practitioners per 100,000 population according to ONS based on health authority of patient residence; NHS facilities per 100,000 population in hospital local health authority; underprivileged area score; one-parent families; mobility. *Other independent variables included in univariate analysis:* Aggregate discharge data: percentage of live discharges who went home; average number of diseased bodily systems; average LOS number of cases. Hospital data: hospital doctors per case; percentage of nurses above an A grade; nurses per doctor and per bed; number of hospital beds; percentage of geriatric beds; location outside of London; non-university teaching; other general hospital; provision of a range of specialist units; hospital income per bed and per case; total and first accident and emergency attendances; hospital character standards; results of survey of patient-centred care. Community attributed data: general practitioners per 100,000 population according to ONS based on individual data averaged at health authority of residence level; general practice nurses per 1000 population according to ONS in hospital local health authority; elderly living alone; children aged under 5; social class V; unemployed; overcrowded accommodation; ethnic minority; percentage of patients with limiting long-standing illness; provision of nursing homes and residential care homes in hospital local health authority area. 3 Three sources 4 Excluded hospitals that had poor-quality data, more than 30% of inpatient episodes without a valid discharge or more than 30% of primary diagnoses recorded as unknown. 5 No – used criteria based on type and size as well as quality of data recorded in the HES database. 6 2 hospitals per health authority across England, 85% of all admissions in England HES database.
Commentary	Only one of the measures of co-morbidity was significant in the model and this might be related to the lack of data on the severity of illness. The figures used were the aggregates for the health authority of hospital location rather than individual figures for each hospital's emergency catchment area (often very different). It could simply be that more doctors mean more admission. The data are an inadequate basis for drawing the overall conclusion that higher hospital doctor and GP ratios equates to lower death rates. The accuracy of the measurement of prognostic or risk factors in patients treated by the hospitals is inadequate. The rates derived from the HES database represent episodes rather than actual patients. Hospital doctors and acute care beds per 100,000 population would be better examined independently. There are other factors outside of the hospital that could affect hospital death rates after the adjustments made in this instance, such as care in the community.
Research implications	Repeat the above analysis with data aggregated by electoral ward of residence rather than by hospital of admission. What about nurses: do they have the same impact on mortality?

ID, origin, authors (year)	812, Netherlands, Tutuarima, J.A., Haam, R.J. and Limburg, M. (1993)
Aims	To examine the relationship between the risk of falling by stroke patients and the number of nurses on the ward in the acute care setting *Workforce:* Nurses, acute-care settings *Feature:* Nursing workload, assessed using the ratio of patients per nurse *Outcome:* Falls by stroke patients
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, case–control with patient interviews 2 Patients admitted within one week after stroke onset 3 349 patients from a cohort of 760 stroke patients sampled for a multi-centre study, 13 neurological wards (some were combined with other specialities) 4 6 months post-stroke 5 Patient data were collected from the medical and nursing charts and senior nursing officers provided ward characteristics.
Results Quantitative results	For both the day and the evening shifts no difference was found between the case and control patients in the composition of nursing staff (the proportion of qualified, trainee or temporary nurses). The night shift showed a significant difference in the patient:nurse ratio (PNR). There was no overall difference between the nursing workload of the case and control patients. The greatest number of falls occurred in the daytime when most nursing staff were present

		Cases			Controls			Difference of	95% CI
Shift	***Falls***	***Patients***	***Nurses***	***PNR***	***Patients***	***Nurses***	***PNR***	**the mean PNR**	
Day	26	25.2	6.8	3.91	25.1	6.6	3.97	–0.06	–0.51/0.39
Evening	17	25.4	3.4	7.99	26.6	3.4	8.23	–0.24	–0.97/0.50
Night	6	28.2	2.2	13.44	26.2	2.3	12.21	1.24	0.28/2.20
All shifts:									
All patients	49	25.6	5.1	6.49	25.8	5.0	6.46	0.04	–0.33/0.40
High-care patients	44	3.6	5.0	0.86	3.4	5.0	0.80	0.06	–0.07/0.19

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Control patients were matched using ward, sex, number of hospital days at the time of the fall, stroke severity and age. Severity was assessed by the Glasgow Coma Scale (GCS) and supplemented with the Allen Scale (I = very severe, II = severe, III = moderate, IV = mild). 2 No 3 Yes 4 21 patients refused to enter the study; 20 patients' relevant data were excluded because of missing values. 23 patients were excluded from the control group after the interviews as they were said to have fallen during their hospital stay but it was not reported on the hospital charts. 5 Convenience sample of 9 of the 23 hospitals from the case–control study 6 Not stated
Commentary	By calculating PNRs this may not take account of the workload faced by the nursing staff. The authors stated that the date of admission did not always coincide with the date of the stroke onset, but all the patients were admitted within one week after the stroke.
Research implications	Are other factors, such as ward organisation, nursing procedures and application of safety procedures more relevant in the prevention of falls? Does the proportion of qualified, trainee or temporary nurses affect the number of falls? Falls are more likely to occur during the day, as patients are more likely to be moving around the ward whereas at night they are likely to be asleep. Looking at different ratios of nursing staff during the day or overall (day and night) would be a better focus of the study.

Table A2.2 Workforce hours to patient ratios

ID, origin, authors (year)	488, USA, American Nurses Association (ANA) (1997)
Aims	To explore implementing nursing's report card: a study of RN staffing, LOS and patient outcomes *Workforce:* Registered Nurses (RNs), secondary care *Feature:* Workforce hours per patient day (total nursing hours per Nursing Intensity Weight (NIW) and RN hours as a percentage of all nursing hours) *Outcome:* LOS and adverse events (pressure ulcers, pneumonia, urinary tract infection (UTI) and post-operative infections)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, cross-sectional 2 Not stated 3 Not stated 4 In-hospital 5 *Hospital Cost Reports:* California: The Annual Hospital Disclosure Report was obtained for all acute hospital facilities; Massachusetts: Hospital Statement for Reimbursement was obtained for all acute hospital facilities; New York: Institutional Cost Reports. *Patient-level data:* California: Office of Statewide Health Planning and Development; Massachusetts: Massachusetts Rate Setting Commission; New York: Statewide Planning and Research Cooperative System. Data collection: 1992 and 1994
Results Quantitative results	

State	Year	Total hours per NIW (%)	% RN hours
Geometric Length of Stay Index			
MA	1992	–9.7	–0.27
MA	1994	Not significant	–0.19
NY	1992	–6.46	–0.19
NY	1994	–4.40	–0.11
CA	1992	–4.82	–0.07
CA	1994	–5.40	–0.16
Pressure ulcer rates			
NY	1992	–17.89	–1.77
NY	1994	Not significant	–1.23
CA	1992	Not significant.	–0.79
CA	1994	–15.59	–1.23

State	Year	Total hours per NIW (%)	% RN hours
Pneumonia rates			
NY	1992	Not significant	Not significant
NY	1994	Not significant	Not significant
CA	1992	Not significant	–0.56
CA	1994	+7.65	–0.39
Postoperative infection rates			
NY	1992	Not significant	Not significant
NY	1994	Not significant	Not significant
CA	1992	Not significant	–0.53
CA	1994	Not significant	–0.47
Urinary tract infection rates			
NY	1992	Not significant	Not significant
NY	1994	Not significant	–0.65
CA	1992	Not significant	–0.64
CA	1994	Not significant	–0.65

The table above shows regression coefficients for all of the outcomes, Massachusetts is not included for some of the analyses due to data quality and lack of statistically significant relationships. They can be interpreted as follows: for New York in 1994 with all other variables held constant (% of RN hours, medical school, other teaching, large urban and rural) an increase of 1 hour of nursing care per NIW would predict a geometric LOS in a hospital 4.4% lower than the average for the state. Overall, more nursing hours per NIW and a higher skill mix of nurses are associated with reduced hospital lengths of stay. In 5 out of 6 cases, total hours per NIW were significantly related to LOS, and in all 6 the RN percentage of total nursing hours was significantly and inversely related to LOS. For pressure ulcers, In 2 of the 4 cases additional hours of nursing per NIW were related to lower rates. Further, in all 4 cases, nursing skill mix was associated with a reduction in pressure ulcers (each additional % of nursing personnel that were registered nurses was associated with a reduction between –0.79 and –1.77). For the other outcomes the results were not as consistent with only California showing a relationship between % RN hours and lower rates throughout, and for New York only 1994 with UTIs was a significant relationship found.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Nursing Intensity Weights are used to recognise differences in patients' acuity of need for nursing care. 2 *Teaching status:* Medical school affiliate: a primary undergraduate medical school; other teaching: other hospitals with at least 40 residents per 10000 discharges; non-teaching: all other hospitals. *Setting:* Large urban: hospitals in Metropolitan Statistical Areas (MSA) with populations over 2,000,000 or in MSAs consolidated into Consolidated Metropolitan Statistical Areas (CMAs) with populations over 2,000,000; urban: hospital in other MSAs; rural: hospitals not located in a MSA. 3 Yes 4 *California:* For 1994, 7 hospitals did not submit cost reports, 26 did not report nursing hours and 8 reported unrealistic nursing hours. For 1992, 12, 25 and 11 respectively. *Massachusetts*: For 1994, 8 hospitals did not submit cost reports, 2 did not report nursing hours and 3 reported unrealistic nursing hours. For 1992, 4, 16 and 3 respectively. *New York:* For 1994, 15 hospitals did not submit cost reports, 27 did not report nursing hours and 21 reported unrealistic nursing hours. For 1992, 27, 30 and 8 respectively. 5 The states were selected because their data are publicly available at a reasonable cost; the data are reasonably current; the states contain a sizeable percentage of the nation's hospitals, patients and nurses; they are representative of any differences in patient care which may be provided in the East compared with the West. The research team also had experience of working with them. 6 California, Massachusetts and New York
Commentary	The quality of the data received from the Hospital Cost Reports was very uneven, especially nursing hours. In each state a large proportion of hospitals had to be excluded from the study due to non-reporting or obvious errant reporting. The reporting of complications in secondary diagnoses was poor. Hospitals natural inclinations will always be to under report such conditions relative to all other diagnoses.
Research implications	Due to the poor quality of the data, if this study was repeated with cleaner and better standardised reporting would relationships be found between staffing levels, skill-mix, LOS and adverse events?

ID, origin, authors (year)	419, USA, Archibald, L.K., Manning, M.L., Bell, L.M. *et al.* (1997)
Aims	To assess the effect of fluctuations in cardiac intensive care unit (CICU) nurse staffing levels and patient census on nosocomial infection rates (NIR) *Workforce:* Registered Nurses, cardiac intensive care unit *Feature:* Workforce hours per patient day (monthly NIR: number of infections per 1000 patient days; the monthly nursing hours per patient day) *Outcome:* nosocomial infections
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, retrospective 2 All patients admitted to the study site during the stated period who experienced an NI 3 782 admissions 4 In-hospital 5 Data on the number of hours worked by CICU nurses each month stratified by the level of training, patient days and NIRs were obtained by review of hospital microbiology, infection control, and patient administrative records during December 1994 through December 1995.
Results Quantitative results	A strong linear relationship was found between monthly NIR and patient days: r = 0.89, p = 0.0001, line slope = 0.065; indicating an increase in the NIR of 6.5 infections per 1000 patient days for each 100-day increase in patient days. An inverse linear relationship was found between monthly NIR and nursing hours to patient day ratio: r = –0.77, p = 0.003, line slope = –1.96, indicating a fall in the NIR of nearly 2 infections per 1000 patient days for each unit increase in the nursing hours-to-patient day ratio. An inverse correlation between NIR and nursing hours was noted when the number of hours worked per month was <7600; when the number exceeded 7600 hours, the correlation became positive; however, the results were not statistically significant. The authors considered that as mediastinal infections could have originated in the operating room they repeated the analysis after excluding this type of infection and similar results were found (NIR and patient days: r = 0.80, p = 0.002; NIR and ratios: r = -0.67, p = 0.02)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 None stated 2 None stated 3 Yes 4 Complete 5 No 6 One children's hospital in Philadelphia
Commentary	The data showed that although NIR initially fell with increased numbers of hours worked by RNs, there is a threshold number of nursing hours per month (~7600 hours in this study) above which the NIR increases. Although patient acuity may be an important factor in determining NIR in the ICU setting it was not found to be a significant risk factor in an initial NIR outbreak study. Although all nurses who provided patient care were trained in paediatric intensive care, the level of experience of individuals and the duration of their experience in the CICU were not determined, thus the two could not be correlated. Stratification of the correlation analysis by organism did not yield statistically significant results because of the small numbers of individual organisms. Quantification in hours of patient exposure to other health workers, such as physicians and physiotherapists, was not feasible as was possible for nurses.
Research implications	Future research in this area should include the number of hours other health care workers actually spend in direct patient care. If health care worker reductions are used as a part of cost containment, the factors that lead to breakdown in infection control as a result of these reductions and the effect of these factors on patient outcomes must further be assessed and preventive measures implemented.

ID, origin, authors (year)	414, USA, Blegen, M.A., Goode, C.J. and Reed, L. (1998)
Aims	To describe, at the level of the nursing care unit, the relationships among total hours of nursing care, Registered Nurse skill mix, and adverse patient outcomes *Workforce:* Nurses, secondary *Feature:* Workforce hours per patient day and skill mix (all hours per patient day = hours of direct patient care (i.e. the employee was assigned to provide care for a patient or group of patients) by RNs, LPNs and nursing assistants each month divided by the patient days of care on the unit for the month; proportion of RN Hours = the hours of direct patient care from RNs divided by patient days divided by all hours per patient day) *Outcome:* Adverse events
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observational 2 5 surgical units, 10 medical, 3 obstetric/gynecology, 8 paediatric, 4 critical care, 4 psychiatric, 2 eye/ear/nose and urology, and 6 orthopaedic and neuroscience units. Ambulatory or outpatient clinics, operating rooms, emergency rooms and delivery rooms were excluded. Psychiatric and other units with low incidence of surveillance were not included in the analyses for infections and decubiti. 3 42 inpatient units, 880-bed hospital, 1074 total FTE (full-time equivalent) nursing staff and 832 of these were RNs. 4 In-hospital 5 All data came from hospital record. Nurse staffing, tenure, and patient days of care each month came from payroll and human resources databases. The quality assurance department provided the data for the adverse events and patient compliant data were obtained from the office of patient relations. Patient acuity data was obtained from files containing monthly acuity system reports. Patient falls and medication error data were gathered from incident reports. Data were used from each month of fiscal year 1993 (July 1992 – June 1993)
Results Quantitative results	The correlations among staffing and outcome variables were investigated and All hours was found to be statistically significantly correlated with infections (0.564**), decubiti (0.573**), complaints (0.427**) and death (0.640**), but not for medication errors (–0.124) or falls (–0.255); RN proportion was correlated with falls (–0.305**), infections (0.158*) and death (0.351**), but not for medication errors (–0.153), decubiti (0.176) or complaints (0.058) *Four multivariate models were evaluated for each dependent variable. Because of collinearity, all hours was excluded in two of the models. In the first regression model for each dependent variable the effects of RN proportion, controlling for patient acuity, were negative for all adverse outcomes except death rates; however, these coefficients were not statistically significant. When all hours of nursing care were added to the analyses in Model 2, the direction of the relationship between RN proportion and the outcomes remained negative and the size increased. The coefficient for complaints became significant. Higher total hours of care were associated with a higher incidence of negative outcomes, but higher RN proportion was related to lower incidence of negative outcomes. Multiple regression modeled the curvilinear relationships among RN proportion and the outcome variables by inserting a dummy variable for the upper 25% of RN proportion. For Model 3, the coefficient for RN proportion increased further for 5 of the 6 variables and became statistically significant for medication errors and decubiti. The relationship between RN proportion and patient falls was small and not significant; however, the coefficient for the dummy variable was negative, unlike the rest. Falls decreased in the upper ranges of RN proportion. For Model 4, the negative relationships between RN proportion and the outcomes remain.*

	Variable	Model	R^2	R^2 adjusted	All hours	RN proportion	Dummy RN >0.875
	Medication error	1 (All hours excluded)	0.030	–0.019		–0.095	
		2 (All hours included)	0.031	–0.045	0.050	–0.105	
		3 (Dummy included)	0.175	0.110		–0.525**	0.556**
		4 (All hours and dummy)	0.186	0.098	–0.202	–0.530**	0.611**
	Falls	1	0.112	0.067		–0.212	
		2	0.112	0.042	–0.019	–0.216	
		3	0.154	0.087		0.018	–0.297
		4	0.161	0.070	0.159	0.021	–0.340
	Infections	1	0.322	0.277		–0.161	
		2	0.377	0.312	0.458	–0.242	
		3	0.344	0.275		–0.330	0.216
		4	0.382	0.294	0.409	–0.325	0.116
	Decubiti	1	0.279	0.231		–0.114	
		2	0.364	0.298	0.571*	–0.216	
		3	0.382	0.318		–0.490*	0.479**
		4	0.421	0.339	0.413	–0.485**	0.379
	Complaints	1	0.179	0.137		–0.225	
		2	0.247	0.188	0.471*	–0.312*	
		3	0.200	0.138		–0.391	0.215
		4	0.251	0.170	0.430	–0.381	0.099
	Death	1	0.352	0.319		0.027	
		2	0.426	0.381	0.491**	–0.063	
		3	0.432	0.388		–0.292	0.413**
		4	0.468	0.410	0.361	–0.284	0.316**
	*p <0.10 and **p <0.05						
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Severity of illness: controlled for using nursing acuity system, 1 to 7 scale 2 Standardising by patient days controlled for the size and occupancy of units 3 Outcomes retrieved from different sources but each outcome collected through the same means 4 Complete 5 No 6 One large university hospital in Iowa						

Commentary	The hours of care excluded administrative or paid, non-worked time such as vacations, sick leave and holidays. To minimise the effects of random fluctuations from month to month the data were aggregated to an annual rate. Although this study controlled for acuity, the indicator may not have been sensitive enough to control for the higher acuity of patients on the units. The results of this study lack generalisability outside of the study unit. Relying on incident reports as the data source for medication errors and falls may be problematic. Although units track these rates as part of their quality improvement monitoring, the rigour with which reports are completed will vary from unit to unit.
Research implications	Multi-institutional studies with standardised and sensitive acuity measures are needed to describe further the relationship between rates of adverse events in units with higher acuity of patient and staff mix. This study needs replicating in other settings with other kinds of hospitals.

ID, origin, authors (year)	517, USA, Blegen, M.A. and Vaughn, T. (1998)
Aims	To determine the relationship between different levels of nurse staffing and adverse events *Workforce:* Nurses, secondary *Feature:* Workforce hours per patient day. All hours of care per patient day (hours of care) and the proportion of those hours of care delivered by RNs (RN proportion). *Outcome:* Medication administration errors, patient falls and cardiopulmonary arrests
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observation 2 Hospitals that were members of a consortium of hospitals that joined together to create the Institute for Quality Healthcare (IQH). Medical, surgical, ICU, obstetric and skilled care units were all included in the study. 3 39 units in 11 hospitals 4 In-hospital 5 Patient occurrence data were extracted from the Comparative Occurrence Reporting Service (CORS) file and were supplemented by the hospital as available. Each hospital submitted information on staffing levels. Data collection: July 1993 to December 1995.
Results Quantitative results	Initial descriptive analyses revealed a statistically significant nonlinear relationship between RN proportion and medication administration errors. Relationships between RN proportion and the other outcome variables were linear. A nonlinear relationship means that the relationship between RN proportion and the rate of medication administration errors differs at different levels of RN proportion. For example, as the RN proportion for the unit increased from 50% to 85%, the rate of medication errors declined; but as the RN proportion increased from 85% to 100% the rate of medication administration errors increased. A dummy variable was created to explore this further. Units with greater hours of care per patient day from all staff had higher rates of medication errors. Units with higher proportions of RN care, up to 85%, had lower rates of medication errors per 10,000 doses; but units with RN proportions >85% had higher rates of medication errors per 10,000 doses. Similar trends were apparent for the effect of RN proportion on medication errors per 1000 patient days. Units with higher proportions of RN care had lower rates of patient falls, but were unaffected for hours of care. Nurse staffing levels were unrelated to cardiopulmonary arrests. **Variables (N)** / **Adjusted R^2** / **Hours of care** / **RN proportion** / **Dummy RN>85** Medication error/doses (199) / 0.4 / 0.497** / –0.576** / 0.483** Medication error/days (276) / 0.27 / 0.323** / –0.278* / 0.248* Falls/days (276) / 0.27 / –0.49 / –0.456** Cardiac arrests/days (207) / 0.48 / –0.95 / –0.080 *0.10 **0.05 using generalised estimation equation
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 To control for the influence of patient severity, the type of patient care unit (medical/surgical, intensive care, obstetrics, and skilled care) and the average severity of patients in the hospital as a whole, as reflected in each hospitals Medicare CaseMix scores for the three years. 2 None stated 3 Yes 4 Not stated 5 No 6 Not stated

Commentary	The data used were routinely collected and therefore limited to certain outcomes. The authors' restricted their selection of hospitals to those hospitals that were members of IQH and therefore these may be unrepresentative of hospitals in general and consequently have biased the results. Adjustments were made for severity but did not use a standard well-established method for it.
Research implications	Why does the complex relationship between the optimum cut-off point of RNs and patient outcomes exist? Further investigation of the nonlinear relationship between RN proportion and medication errors is needed. What is it about nursing procedures regarding medication administration on units with high proportions of RNs that results in higher rates of medication administration errors? Is it that with higher RN proportions there is a heightened vigilance and therefore more reporting? Units with higher RN proportions have more severely ill patients who need more complex medications and therefore there are more opportunities for error? Units with higher RN proportions have less total personnel than needed for optimal patient outcomes?

ID, origin, authors (year)	175, USA, Bond, C.A., Raehl, C.L. and Franke, T. (2001)
Aims	To evaluate hospital demographics, staffing, pharmacy variables, health care outcome measures and medication errors *Workforce:* All staff, secondary care *Feature:* Workforce hours per number of occupied beds and skill mix (ratios of staff) *Outcome:* Medication errors
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, survey 2 Responders from the 1992 database, only full-time personnel 3 1116 hospitals, 430,586 medication errors 4 In-hospital 5 Hospital medication error information was collected as part of the 1992 National Clinical Pharmacy Services database survey. Pharmacy directors were asked whether their hospital had a medication error reporting system, the total number of medication errors for the previous 12 months and the number of medication errors determined to adversely affect patient outcomes. Data for pharmacy teaching affiliation, pharmacy directors degree, pharmacists location within each hospital and drug costs were obtained from the 1992 database of the National Clinical Pharmacy Services. Mortality rate information was obtained from the Health Care Financing Administration. Data on census region information, size, hospital ownership, hospital staffing, admissions data, occupancy rates, teaching affiliation, length of stay and total cost of care for each hospital were obtained from the American Hospital Association's (AHA) Abridged Guide to the Health Care Field.
Results Quantitative results	Simple regression analysis showed that as the number of hospital administrators, Registered Nurses, ratio of RNs to licensed practical/vocational nurses, registered pharmacists, physical therapists, and total personnel increased then number of medication errors also increased. Conversely, as the number of medical residents and pharmacy technicians increased the number of medication errors decreased. The only variable that had a statistically significant association with the number of medication errors that adversely affected patient outcomes was the number of medical residents per occupied bed.

Hospital personnel	**Mean no. staff per 100 occupied beds**	**Slope**	**SE**	**Significance**	**95% CI**
Administrators	6.98 ± 9.07	2.9931	1.516	0.04	0.1814, 5.968
Physicians	35.41 ± 18.37	–0.8384	0.6359	not significant	–2.0852, 0.4092
Ratio of board certified physicians to all physicians	68.89%	0.036	0.0661	not significant	–0.9818, 0.1614
Medical residents	5.12 ± 15.26	–3.1541	1.067	0.0032	–5.2492, -1.0589
Registered Nurses	112.67 ± 65.73	0.6908	0.263	0.0008	0.2860, 1.0956
Licensed practical/vocational nurses	29.79 ± 31.03	–0.0045	0.0075	not significant	–0.0193, 0.0103
Ratio of registered nurses to licensed practical/vocational nurses	3.19 ± 4.98	2.5563	0.8914	0.0314	0.7619, 4.2971
Physician assistants	0.32 ± 1.34	–0.0755	0.0506	not significant	–0.0237, 0.2938
Registered pharmacists	7.21 ± 4.04	9.996	4.2882	0.002	1.5624, 18.4315
Pharmacy technicians	5.81 ± 3.89	–0.0529	0.0177	0.0029	–0.0876, -0.0181
Ratio of registered pharmacists to technicians	1.24 ± 1.36	0.0097	0.1146	not significant	–0.2165, 0.2470
Medical technologists	13.57 ± 9.54	3.9998	1.4722	not significant	1.1108, 6.8888
Dieticians	1.8 ± 1.95	10.7244	8.2919	not significant	–5.5469, 26.9958
Occupational therapists	1.25 ± 3.05	2.5787	7.6464	not significant	–12.4258, 17.5834
Physical therapists	3.26 ± 4.28	21.4581	4.6502	0.0001	12.3329, 30.5833
Respiratory therapists	5.98 ± 5.02	4.5336	2.9481	not significant	–1.2515, 10.3189
Social workers	2.97 ± 3.06	2.2783	5.1357	not significant	–7.8049, 12.3506
Total	506.32 ± 284.23	0.1128	0.0456	0.0082	0.0292, 0.1963

	The next table shows a multiple regression analysis but only statistically significant associations are reported. Factors associated with increased total medication errors per occupied bed per year were number of registered nurses per occupied bed and number of registered pharmacists per occupied bed. A decrease in medication errors was associated with the number of medical residents per occupied bed. **Variable** / **Slope** / **SE** / **Significance** / **95 % CI** Medical residents / –1.478 / 0.5251 / 0.0014 / –2.3601, –0.3448 Registered nurses / 1.624 / 0.758 / 0.032 / 0.1361, 3.1119 Pharmacists / 25.0573 / 7.71461 / 0.0001 / 11.0199, 39.0948
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 None for the simple regression, but for multiple regression the severity of illness (percentage of ICU days, annual number of emergency room visits divided by average daily census, and % of Medicaid patients) was used. 2 All of the following variables were included in the regression analysis: *Size:* small, medium or large; *Hospital pharmacy teaching affiliation:* affiliation with college of pharmacy, no college of pharmacy affiliation but an affiliation with other health education programmes, or no affiliation with any education programme; *Hospital teaching affiliation:* Teaching or non-teaching; *Education:* BS, PharmD, MSPharmacy, MBA, PhD, or non-pharmacy masters; *Hospital Ownership:* Non-federal government, non-profit, for-profit and Federal Government; *Pharmacist's predominant location*: decentralised, centralised with ward visits or centralised. 3 Yes 4 Possible 3444 hospitals, only 1597 (46%) responded and 1116 provided the correct information. 5 No 6 Regions: New England, Mid-Atlantic, South Atlantic, East North Central, East South Central, West North Central, West South Central, Mountain and Pacific
Commentary	Data from this study are from 1992 and may not be representative of health care in 2001. It is possible that the information provided to the authors was inaccurate, as the results were not verified. The hospitals in the study population may not be representative of all hospitals in the USA, but did represent 32% of all US hospitals. Given that this was a population-based survey study, the authors could not determine the specific information about each medication error and the types of harm experienced by patients. Since medication errors were likely to be underreported, actual error rates were likely to be higher than reported.
Research implications	Further study is needed to determine the specific reasons why medication errors are affected by hospital size. Specific exploration of actual workloads of the workforce in relation to medication errors is needed. Does a highly educated and trained workforce reduce medication errors?

ID, origin, authors (year)	333, USA, Cho, S.H., Ketefian, S., Barkauskas, V. *et al.* (2003)
Aims	To examine the effects of nurse staffing on adverse events, morbidity, mortality and medical costs *Workforce:* Nurses, acute-care hospitals *Feature:* Workforce hours per patient day (all hours: total productive hours worked by all nursing personnel per patient day; RN Hours: total productive hours by Registered Nurses per patient day; RN proportion: RN hours divided by all hours) *Outcome:* Adverse events (only when not present at admission) Morbidity (measured indirectly by LOS), mortality and costs were also investigated but not in relation to staffing levels, hence will not be reported here.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, retrospective 2 Excluded: government, long-term care, and 'non-comparable' hospitals defined by Office of Statewide Health Planning and Development (OSHPD) and medical patients. 20 common surgical diagnostic related groups (DRGs) were selected as the patient groups. 3 232 hospitals; 124,204 patients 4 In-hospital 5 Two databases: Hospital Financing Data produced by California's OSHPD were used for nurse staffing and hospital characteristics from three fiscal years (1996/97, 1997/98 and 1998/99); State Inpatient Databases (SID) California–1997 released by the Agency for Healthcare Research and Quality (AHRQ) included information of inpatients who were discharged from the Californian hospitals for 1997.
Results Quantitative results	RN hours and RN proportion had a significant inverse relationship with pneumonia and all RN hours had a positive relationship with pressure ulcers. All other relationships were not statistically significant. **Outcome** / **All hours OR (95% CI)** / **RN hours** / **RN proportion** Patient fall/injury / 1.08 (0.99–1.18) / 1.07 (0.96–1.19) / 0.96 (0.21–4.49) Pressure ulcer / 1.13 (1.01–1.27)* / 1.11 (0.97–1.27) / 0.75 (0.11–4.98) Adverse drug event / 1.04 (0.96–1.13) / 1.01 (0.92–1.11) / 0.62 (0.16–2.38) Pneumonia / 0.96 (0.91–1.13) / 0.91 (0.85–0.97)** / 0.37 (0.15–0.91)* Urinary tract infection (UTI) / 1.02 (0.95–1.08) / 1.01 (0.93–1.08) / 0.92 (0.31–2.64) Wound infection / 1.00 (0.95–1.06) / 0.97 (0.91–1.04) / 0.52 (0.21–1.30) Sepsis / 1.01 (0.95–1.08) / 1.02 (0.95–1.09) / 1.20 (0.43–3.33) *p <0.05 and **p <0.01
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 *Patient characteristics:* Age, sex, ethnicity, primary payer, DRG, number of diagnoses at admission (to reflect severity and co-morbidity) and type of admission (scheduled or unscheduled) 2 *Hospital characteristics:* Ownership (non-profit or investor owned), hospital size (small = 1–99 beds, medium = 100–299 and large = 300+), teaching affiliation (teaching or non-teaching) and location (rural or non-rural) 3 Yes 4 Complete 5 No 6 Statewide: California

Commentary	A major finding in this study was the great impact of patient characteristics on the occurrence of adverse events, while hospital characteristics had minimal influence. This study reported lower adverse event rates compared to previous studies. This could have been because the authors excluded diagnoses present at admission, which did lower the incidence rate. The use of ICD-9 codes to detect adverse events may have caused underreporting and consequently lower incidence rates than actually occurred. Another reason is that medical patients were excluded, who are likely to be more seriously ill than surgical patients. Aggregated nurse staffing measures may have smoothed the level of staffing over the year, thus did not account for the variability in either patient census or in nursing hours. This study focused on quantifying nurse staffing levels, while professional characteristics of nursing personnel (e.g. experience, educational preparation, and certification) that may also influence patient outcomes were not considered. Organisational characteristics of the hospitals were also not investigated.
Research implications	What is it that RNs have or do that produces better outcomes? Is it because they have a higher level of knowledge and skill? Further study is needed to add risk factors specific to pressure ulcers in surgical patients, such as immobility, malnutrition, operating time and conditions on the operating table, to isolate the effects of nurse staffing on pressure ulcers form those of patient risk factors. Future studies need to evaluate the appropriateness of ICD-9-CM codes in examining nursing care quality

ID, origin, authors (year)	1102, USA, Kovner, C. and Gergen, P.J. (1998)
Aims	To examine the relationship between nurse staffing and selected adverse events hypothesised to be sensitive to nursing care *Workforce:* Registered Nurses, secondary care *Feature:* Staffing levels (Calculated as the number of full-time equivalent (FTE) RNs working in the hospital and outpatient departments per adjusted patient day*) *Outcomes:* Adverse events *Nurse-sensitive:* Venous thrombosis or pulmonary embolism after major surgery or vascular procedure excluding discharged patients with venous thrombosis as principal diagnoses; nosocomial urinary tract infections (UTIs) after major surgery excluding discharged patients in MDC 11 (Renal), MDC12 (Male genital) or MDC 13 (Female genital); pneumonia after major surgery or an invasive vascular procedure excluding discharged patients in MDC 4 (Respiratory), with cancer, or with AIDS. *Non-nurse-sensitive:* Pulmonary compromise after major surgery (pulmonary congestion, lung oedema, or respiratory insufficiency or failure), excluding discharged patients in MDC 4 or MDC 5 (Cardiovascular); AMI after major surgery excluding discharged patients in MDC 5; gastrointestinal haemorrhage or ulceration after major surgery excluding discharged patients in MDC 6 (Gastrointestinal) or MDC 7 (Hepatobiliary); mechanical complications because of device, implant or graft, excluding organ transplant.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Non-experimental, survey 2 All discharged patients 18 years+ were included 3 506 acute-care hospitals 4 In-hospital 5 Hospital characteristics data were collected in 1993 from the American Hospital Association (AHA) (e.g. staffing, beds and services provided) and matched by hospital identification number to discharge data from the Nationwide Inpatient Sample (NIS) from the Agency for Health Care Policy and Research (AHCPR).
Results Quantitative results	A significant relationship existed between nurse staffing and one of the non-nurse-sensitive outcomes. An inverse relationship existed between FTE RNs per adjusted inpatient day and UTI infections (p <0.001) and pneumonia (p <0.01), between RNAPD and thrombosis (p <0.01) and pulmonary compromise (p <0.05). An increase of 0.5 RN hours/patient days is associated with 16% decrease in UTI, a 4.2% decrease in pneumonia, a 2.6% decrease in thrombosis and 1.8% decrease in pulmonary compromise after surgery. No relationship was found between hospitals with FTE nurse practitioners (NPs) and the adverse events, but those with more physician assistants had higher rates of pneumonia and thrombosis after surgery.

Predictors	**Thrombosis after major surgery**	**UTI after major surgery**	**Pneumonia after major surgery**	**Pulmonary compromise after major surgery**
FTE RNs to adjusted patient days	–33.22 (–57.76,-8.68)	–639.96 (–852.78, –421.15)	–159.41 (–252.67, –66.16)	–56.96 (–117.62, –1.76)
FTE nurse practitioner	p-value = 0.35	p-value = 0.14	p-value = 0.07	p-value = 0.86
FTE physician assistant	0.008 (0.002, 0.01)	p-value = 0.06	0.02 (0.01, 0.03)	p-value = 0.42
N	478	470	476	478

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Medicare Case Mix Index and proportion of patients using Medicare were used to adjust for case mix. 2 Adjustment was also made for hospitals that were urban or rural, teaching status (AMA-approved residency programme, or membership in the Council of Teaching Hospitals (COTH), ownership (government, private not-for-profit, and private investor-owned), bed size, hospital resources (has a nursing school, the number of FTE NPs employed, the number of FTE physician assistants employed, and the non-RN staff employed (defined as total FTE hospital employees minus total FTE RNs)), region's and hospital's relationship with a managed-care organisation (participant in a network, affiliated with an HMO, affiliated with a PPO, HMO product owned or provided through other formal arrangement, and PPO product owned or provided through other formal arrangement). 3 Yes 4 Six states excluded because they didn't include the day on which the principal procedure was performed and a seventh did not permit linking AHA to NIS data. 83 hospitals were excluded from the analysis because they were not in operation for a full year (n=50), children's hospitals (n=6), Medicare case-mix data were unavailable (n=21), and outliers (extremely high or low staffing levels n=6). 5 An AHCPR representative selected a 20% stratified probability sample of hospitals to approximate US community hospitals for 1993. The sample was weighted so either nationwide or state-specific estimates could be made. 6 Hospitals from 10 (California, Colorado, Connecticut, Florida, Maryland, Massachusetts, New Jersey, New York, Pennsylvania, Wisconsin) of the 17 states in the NIS.
Commentary	This study had a methodological aspect – testing the relationship between nurse staffing and nurse-sensitive and non-nurse-sensitive adverse events. Nurse-sensitive outcomes – those that can be directly linked to care given by a nurse, e.g. nosocomial urinary tract infections (UTIs) because nurses are responsible for urinary catheter care. Hospitals with high RNAPD also tended to have a high ratio of FTE non-RNs per adjusted patient day (non-RNAPD). It is therefore possible that high overall staffing levels are inversely related to adverse events, rather than just RNs. Post-hoc analyses in which skill mix without RNAPD was an independent variable in the model were conducted and skill mix was found to be inversely related to pneumonia after surgery (–1.2, p <0.004). Due to coding inconsistencies in the Healthcare Cost and Utilisation Project Quality Indicators (HCUP QI) it may not accurately assess surgically related outcomes and the discharge abstracts may not be correct. The selection of the sample may have biased the results. Patients eliminated because of multiple diagnoses might respond differently to nurse staffing levels from those with relatively simple diagnoses. Patient characteristics could not be controlled for due to the data sets and could be associated with the outcomes. Using nursing home days in the denominator may have biased the results. The r-squares in the model were low (7.7%, 11% and 25.2%) and thus anywhere from 77% to 92% of the variance remained unexplained.
Research implications	Further research is needed to specify better models in this area. Other studies are needed to examine these variables using more recent data and from different geographical areas.

* Adjusted patient day included the number of patient days in hospital, the number of patient days in the hospital's nursing home, and an adjustment for the number of outpatient visits that reflected the percentage of the hospital budget devoted to the outpatient departments as a part of the total facility's budget. Thus the ratio of RNs to adjusted patient day does not make hospitals with large outpatient departments appear to have higher staffing levels than those facilities without outpatient departments.

ID, origin, authors (year)	643, USA, Kovner, C., Jones, C., Zhan, C. *et al.* (2002)
Aims	To examine the impact of nurse staffing on selected adverse events hypothesised to be sensitive to nursing care between 1990 and 1996, after controlling for hospital characteristics *Workforce:* Registered Nurses (RNs), Licensed Practical Nurses (LPNs), physicians and dentists, residents and interns, secondary care *Feature:* Workforce hours per patient day (number of FTE RNs and LPNs working in the hospital and outpatient department per adjusted patient day; hours paid are reported, not hours worked) *Outcome:* Adverse events
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observation 2 All non-maternal/non-neonatal discharges age 18+ with major surgery procedure on day 1 or 2 of admission. Patients were limited to those who were admitted from the emergency room or as a planned admission thus eliminating patients admitted from nursing homes, other hospitals and elsewhere. 3 530–570 hospitals for each of the years from 1990–1996, with 187 hospitals having data for all 7 years 4 In-hospital 5 Nurse staffing data from 1990–1996 were obtained from the American Hospital Association (AHA) annual survey of hospitals; adverse event data were obtained from the National Inpatient Sample (NIS) for the same period.
Results Quantitative results	After controlling for other variables RN hours per adjusted patient day were inversely related to all adverse events, but was significant (p <0.05) only for pneumonia. The LPN hours per adjusted patient day were not significantly associated with any adverse events. Between the other staffing variables, resident/intern hours per adjusted patient day were positively (p <0.05) related to all adverse rates except UTI.

Explanatory variable	**Thrombosis**	**Pulmonary compromise**	**UTI**	**Pneumonia**
RN Hours per adjusted patient day	–0.0002 (0.0082)	–0.0047 (0.0074)	–0.0064 (0.0055)	–0.0169 (0.0077)*
LPN Hours per adjusted patient day	–0.0399 (0.0260)	0.0023 (0.0221)	0.0065 (0.0157)	–0.0080 (0.0251)
MD/DDS Hours per adjusted patient day	–0.0664 (0.0299)*	–0.0116 (0.0270)	–0.0325 (0.0192)	0.0098 (0.0281)
Resident/intern hours per adjusted patient day	0.1004 (0.0294)**	0.0382 (0.0168)*	0.0009 (0.0114)	0.0427 (0.0177)*

*p <0.05 and **p <0.01

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Medicare Case Mix Index for each year, proportion of patients for whom Medicare was the principal payer, proportion of patients for whom Medicaid was the principal payer and source of admission were used to adjust for severity. In addition a year-specific, fixed effect in the regression was used as an additional control for case mix. 2 Location (urban or rural), teaching status (membership in the Council of Teaching Hospitals (COTH)), ownership (government, private not-for-profit, and private investor-owned), bed size, region, hospital affiliation with HMO or PPO, and hospital-owned nursing school. 3 Yes 4 Some states and hospitals were excluded because: the state did not require hospitals to report the day on which the principal procedure was performed as part of the discharge data; the state did not permit linking of NIS data to the AHA database; hospitals were not in operation for a full calendar year; and/or hospitals were exclusively children's hospitals. 5 No 6 Six states for 1990–1992, four additional states for 1993–1994, and three more states for 1995–1996 (total 13 states)

Commentary	AHA data do not distinguish between direct-care RNs and those RNs employed by the hospital in indirect or management roles. Unfortunately, if the increase in RN staffing was for RN managers this could blunt any impact of staffing increases on patient outcomes. Moreover, the AHA staffing data reflect paid hours and therefore are likely to overestimate productive hours. Another limitation is that the AHA data set does not include unlicensed assistive personnel. The finding that resident intern hours were positively related to adverse events could reflect problems that occur when residents rather than more experienced physicians are responsible for care. On the other hand it could reflect that residents work in facilities with more severely ill patients and that the case mix adjustments used in this study did not account for the severity. The study may have lacked statistical power to identify the independent effect of nurse staffing. HCUP QIs are indicators of quality, but they are subject to many sources of errors inherent in administrative or claims data.
Research implications	Research is needed to investigate how nurse staffing actually affects quality in general and how nurse staffing interacts with other factors, such as physician staffing, hospital beds, etc., in determining quality. More work is needed to understand staffing mix relative to patient groups, acuity and the ultimate impact on quality. The optimal level of nurse staffing needed to produce high-quality, cost-effective patient care remains largely unknown. More accurate and consistent measures of acuity and quality, and more complete data on nurse staffing across all levels of nursing staff are needed in future studies to more clearly explain the complex relationship between staffing and quality of care.

ID, origin, authors (year)	385, USA, Robertson, R.H. and Hassan, M. (1999)
Aims	This study addresses the questions of whether skill mix and staffing intensity in non-physician caregiver groups have an effect on risk-adjusted mortality, and whether the quality of care provided to patients with COPD is sensitive to the staffing intensity and skill mix within specific caregiver groups. *Hypothesis 1:* Hospitals with higher staffing intensities of nurses and ancillary nurses will have lower risk-adjusted mortality rates for patients with COPD. *Hypothesis 2:* Hospitals with higher staffing intensities of respiratory care practitioners will have lower risk-adjusted mortality rates for patients with COPD. *Hypothesis 3:* Higher staffing intensities of RADG and other radiologic workers will have lower risk-adjusted mortality rates for patients with COPD. *Hypothesis 4:* Hospitals with higher staffing intensities of laboratory technologists and other laboratory personnel will have lower risk-adjusted mortality rates for patients with COPD. *Hypothesis 5:* Hospitals with higher skill mixes will have lower risk-adjusted mortality rates for patients with COPD. *Workforce:* Administrators and assistant administrators (ADMIN), physicians (MD), medical residents and interns (RES), Registered Nurses (RN), licensed practical and vocational nurses (LPN), ancillary nursing personnel (ANNUR), respiratory therapists (RESPTH), respiratory therapy technicians (RESPTE), radiographers and radiologic technologists (RADG), radiation therapists (RADT), nuclear medicine technologists (NUCM), other radiologic personnel (RADO), medical technologists (MEDT), other laboratory personnel (LABO), pharmacists (PHAM), occupational therapists (OT), physical therapists (PT) and dieticians (DIET) *Feature:* Staffing levels (number of FTE personnel employed within each group per 100 adjusted admissions) and skill mix *Outcome:* Mortality of COPD patients
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, observational 2 All hospitals reporting data in the AHA's Annual Survey and having mortality data reported in the HCFA Hospital Information Reports. Hospitals that treated enough patients with a primary diagnosis of COPD to have had a predicted 30-day post-admission mortality over 5 were included in the study. 3 Not stated 4 In-hospital 5 Hospital characteristics and staffing data were obtained from the American Hospital Association (AHA) 1989, 1990 and 1991 Annual Survey Data (AHA, 1990, 1991 and 1992). The observed and predicted 30-day post-admission mortality was obtained from the Health Care Financing Administration (HCFA's) Hospital Information Reports for 1989, 1990 and 1991 (HCFA, 1992, 1993). Data collection: 1989–1992.

Results
Quantitative results

Although the thee models are statistically significant, the R-squared values are low. The models only explain 4.9%, 4.7% and 2.9% of the variation in the natural log of the risk-adjusted mortality. Hypotheses 1, 3 and 4 are not supported by the results. Hypothesis 2 is supported by the results. The results for hypothesis 5 are inconclusive. In general, the overall results for the regression controlling for the previous year remain unchanged. The only group of health workers for whom increasing staff intensity is consistently associated with improving outcomes is respiratory care practitioner.

RESPTE	–1.498 (–3.495)**	–0.318 (–1.012)	–1.076 (–3.995)**	–0.742 (–2.219)*	–1.027 (–3.710)**
RADG	0.459 (1.637)	0.440 (1.700)	0.128 (0.601)	0.030 (1.114)	–0.126 (–0.568)
RADT	–0.329 (–490)	1.056 (1.469)	0.704 (1.440)	0.878 (1.195)	0.274 (0.480)
NUCM	–2.274 (–2.068)	–1.807 (–1759)	–0.607 (–0.905)	–1.729 (–1.607)	–0.867 (–1.296)
RADO	0.471 (1.758)	0.223 (0.990)	–1.77 (–0.872)	–0.013 (–0.057)	–0.273 (–1.369)
MEDT	–0.148 (–769)	0.039 (0.229)	0.145 (1.021)	–0.110 (–0.611)	0.190 (1.293)
LABO	–0.267 (–1.451)	–0.372 (2.299)*	0.187 (1.287)	–0.341 (–2.044)*	0.176 (1.232)
PHAM	0.716 (1.268)	–0.725 (–1.433)	0.018 (0.041)	–1.038 (–1.855)	0.0031 (0.007)
OT	–0.694 (–0.863)	0.205 (0.300)	0.653 (0.992)	–0.233 (–0.309)	0.233 (0.331)
PT	0.210 (0.402)	0.516 (1.138)	0.298 (0.738)	0.572 (1.153)	0.461 (1.101)
DIET	–0.057 (–0.096)	1.589 (2.821)**	0.168 (0.254)	1.744 (2.26)*	0.161 (0.217)
R-Squared	0.049	0.047	0.029	0.114	0.096
F	3.027	3.107	2.012	6.49	6.024

*p <=0.05 and ** p <=0.01; t-values in parentheses

Quality appraisal
1 Case mix adjustment
2 Other adjustment
3 Uniform data collection
4 Participant follow-up
5 Random sampling
6 Geographical dispersal

1 *Severity of illness:* % of inpatient days in special care, % of inpatient days paid by Medicaid, Medicare Case Mix Index, emergency room visits / average daily census ratio; Medicare Case Mix Index for hospitals was purchased from the Commission on Professional and Hospital Activities and HCIA Inc.
2 *Financial status*: Average patient occupancy, total expense per adjusted admission, location in a metropolitan statistical area, membership in the Council of Teaching Hospitals, presence of a written contract with a Health Maintenance Organisation;
Ownership: For-profit, government or not-for-profit
Technological sophistication: % of high technology services offered
Size: Beds set up and staffed
3 Yes
4 Not stated
5 No
6 Not stated

Commentary

Some hospital characteristics that could influence the quality of outcomes could not be measured directly (e.g. experience of staff and standard practices used by workforce in the treatment of certain conditions). Adjustment for the experience of the staff and the hospital surgical volume, which have been shown in other studies to be important variables, were not made. The broad measure of severity used in this study may have not adequately controlled for the severity of the illness within the specific diagnostic group.

Research implications

Do respiratory care practitioners possess more skills and more highly refined skills needed in this patient population in order to prevent adverse events leading to death, to detect adverse events early before they become irreversible, and to respond effectively in the event of adverse events?
Outcome measures other than mortality should be used to monitor the impact of the workforce on the quality of care.
Are outcomes in other diagnostic groups sensitive to staffing levels of specific caregivers?
Future studies could also be useful in clarifying the effect of skill mix within caregiver groups.
Experimental studies in this area would be helpful to improve the inference of causality.

ID, origin, authors (year)	140, USA, Sovie, M., Jawad, A.F. (2001)
Aims	To describe restructuring in the organisation and delivery of patient care and the effects of nursing structure and processes on selected patient outcomes *Workforce:* Registered nurses (RN), secondary care *Feature:* Workforce hours per patient day. The hours worked per patient day (HWPPD) for all staff and for RNs (RNHWPPD) *Outcome:* Fall rate, nosocomial pressure ulcer and urinary
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, retrospective cohort 2 Included university hospitals that had more than 300 acute operating beds and adult acutely ill patients hospitalised 3 29 hospital nurses 4 3 years; In-hospital 5 Structure data were collected from the MECON-PEERx Operations Benchmarking Database Reports: full-time equivalents for each type of nursing staff; skill mix; hours worked per patient day (HWPPD) for all staff, and for selected categories of staff; labour costs per discharge. Process data were from the Management Practices and Organizational Processes Questionnaire (MPOP); and the Quality of Employment Survey. Outcome data were from hospitalised patients. Patient satisfaction data for 16 hospitals were from Picker Institute Patient Satisfaction Survey, and 7 hospitals used the Press, Ganey Patient Satisfaction Survey.
Results Quantitative results	The following table shows the significant results (p-value ≤0.05) of the regression analyses using the model: patient outcome = constant + B1 (structure variable) + B2 (process variable). R^2's are the coefficients of determination for the regression models.

Models using RNHWPPD as the structure variable

Patient outcome	**Year**	**Coefficient for RNHWPPD (standard error)**	**Process variable**	**Coefficient for the process variable (standard error)**	**R^2**
Fall rate	1998	–0.51 (0.18)	Unit medical leadership	2.58 (0.93)	0.38
	1998	–0.60 (0.19)	Communication between nurses and physicians	2.87 (1.14)	0.35
	1998	–0.43 (0.18)	Collaboration between nurses and physicians	1.98 (0.92)	0.31
	1998	–0.48 (0.18)	Conflict resolution between the nurses and physicians	2.73 (1.05)	0.36
	1998	–0.49 (0.18)	Nurse decision making	–1.50 (0.68)	0.32
Patient satisfaction	1998	2.87 (1.30)	Inter-unit work relates	17.17 (3.44)	0.67

	Models using HWPPD as the structure variable **Patient outcome** / **Year** / **Coefficient for HWPPD (standard error)** / **Process variable** / **Coefficient for the process variable (standard error)** / **R^2** **Medical units** UTI rate / 1997 / –0.50 (0.23) / Collaboration between nurses / 2.63 (1.27) / 0.36 1997 / –0.54 (0.23) / Collaboration between nurses and physicians / 4.27 (1.90) / 0.38 1997 / –0.65 (0.23) / Nurse autonomy / –3.72 (1.72) / 0.37 Patient satisfaction with pain management / 1997 / –2.30 (1.04) / Unit medical leader / 26.51 (9.01) / 0.40 1997 / –2.45 (1.15) / Information exchange between nurses and physicians / 21.92 (8.77) / 0.33 1997 / –2.28 (1.01) / Communication between nurses and physicians / 19.41 (6.21) / 0.42 Patient satisfaction with education / 1997 / –2.25 (0.95) / Information exchange between nurses and physicians / 24.16 (7.01) / 0.46 **Surgical units** Fall rate / 1997 / –0.33 (0.14) / Conflict resolution between the nurses and physicians / –1.98 (0.73) / 0.34 Nosocomial pressure ulcers / 1998 / –0.32 (0.15) / Unit nurse manager / –2.10 (0.79) / 0.28 Patient satisfaction with pain management / 1998 / –1.40 (0.32) / Nurse decision making / –3.32 (1.14) / 0.57 1998 / –1.21 (0.34) / Achieving patient outcomes / 3.06 (1.46) / 0.49 Increased RN hours worked per patient per day were associated with lower fall rates and higher patient satisfaction levels.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, admission Total Dependence Score (TDS) 2 Standardising the hours worked by patient day controlled for occupancy and size. 3 Uniform, each hospital validated its own submitted data. 4 Complete 5 No 6 One state

Commentary	Process data were collected exclusively from registered nurses who provided patient care. UAP and others who provided patient care should have been included in the data collection. Patient satisfaction data were collected from different survey instruments that used different scales. Though effort was made to standardise the scores, the weaknesses in this methodology were acknowledged. Case Mix Index was the only risk adjuster that was collected; this was available for hospital-level data only. Consequently, there were no risk adjustments made in unit-level data.
Research implications	The findings do not enable predictions of specific ratios or hours by category of staff that result in the best outcomes for patients; more definite answers regarding these structure variables await further research. The RN percentage is a structure variable that by itself provides no helpful information regarding staffing levels on patient care units. It is essential to have HWPPD and RN HWPPD. It is RNHWPPD that can inform patients, staff and the community about the amount of professional nursing care. An optimal balance of RNHWPPD and UAPHWPPS must be achieved to assure quality outcomes at controlled costs. The value in patient care, defined as the relationship between quality and costs, does not come in one size. No single staffing pattern resulted in best value. Patterns for value were tailor made for each institution/unit. Organisation and unit cultures, nursing and medical leadership, collaborative relationship with physicians and other staff, sufficient number of nurses and assistive staff, and adequate support services interact to produce desired outcomes at controlled costs. Hospitals and their nursing departments can agree on important variables and standardised definitions, and can collect data systematically to evaluate care. Select structure and outcome data at the unit and hospital level should be collected annually and reported as a part of the required hospital data that are submitted to state and federal funding agencies. It should be mandatory that these data are available to health care consumers. Reportable data elements should include HWPPD, RNHWPPD, and outcome data. Case Mix Index (CMI) is currently calculated at the hospital level, it also should be calculated at the unit level for risk adjustment purposes. The rigid staffing regulations would be best deferred until the recommended structure and outcome data from all acute care hospitals are systematically collected, reported, and analysed. The necessary and sufficient research findings are not available for such regulations to be evidence-based.

ID, origin, authors (year)	72, USA, Whitman, G.R., Yookyung, K., Davidson, L.J. *et al.* (2002)
Aims	To determine the relationships between nursing staffing and specific nurse-sensitive outcomes across specialty units. *Workforce:* Nurses, specialty units (cardiac and non-cardiac intensive care, cardiac and non-cardiac intermediate care and medical–surgical) *Feature:* Workforce hours per patient day (included total worked hours (paid hours minus sick, vacation and holiday hours) for all staff (RN, licensed practical nurses, nursing aides and secretaries; worked hours per patient day (WHPPD) = total worked hours / monthly patient days for each unit) *Outcome:* Central line infections (CLI), pressure ulcers, medication errors and falls
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, secondary analysis of observational data 2 Cardiac ICU (CICU, 15), non-cardiac ICU (NCICU, 7), cardiac intermediate care (CIMC, 18), non-cardiac intermediate care (NCIMC, 12) and medical–surgical (MS, 43). Obstetric, psychiatric and paediatric units were excluded. 3 95 patient care units across 10 adult care hospitals 4 In-hospital 5 Infection control staff or their designees conducted monthly surveillance for CLI rates. A system-wide, one-day prevalence study was conducted monthly of all patients on all units to obtain pressure ulcer data. Medication errors and falls data were retrieved from reports provided to the risk management offices within the hospitals. Data collection: 1 January 1 to 31 December 1999.
Results Quantitative results	No statistically significant relationships were found between the outcomes of CLI and pressure ulcer rates and WHPPD across the specialty units. An inverse relationship between WHPPD and falls was present in CIMC. Medication error rates were inversely related to WHPPD in the CICU and NCIMC.

Outcome	**Worked hours per patient day**				
	NCICU	***CICU***	***NCIMC***	***CIMC***	***MS***
CLI	not significant	not significant	not significant	not significant	not significant
Pressure ulcer	not significant	not significant	not significant	not significant	not significant
Fall	not significant	not significant	not significant	–0.53 (p <0.05)	not significant
Medication error	not significant	–5.5 (p <0.05)	–0.65 (p <0.05)	not significant	not significant

The worked hours per patient day were (mean, SD):
NCICU = 18.9, 1.4; CICU = 18.8, 4.1; NCIMC = 8.9, 2.8; CIMC = 8.4; 0.9 and MS = 4.0, 1.2.

Quality appraisal	
1 Case mix adjustment	1 None stated
2 Other adjustment	2 Standardising the hours worked by patient day controlled for occupancy and size.
3 Uniform data collection	3 Different methods of data collection were used for different outcomes.
4 Participant follow-up	4 Not stated
5 Random sampling	5 No
6 Geographical dispersal	6 Not stated

Commentary	Included hours of indirect care as a measure of staffing levels, e.g. secretary and nurse managers. These findings suggest that environments with higher-acuity patients and, most likely, more numerous and more complex medication regimes per patient are sensitive to staffing alterations. Risk adjustment was only applied through specialty classification, because the hospitals in the study did not employ a common patient classification or acuity system. The classification of the ICU and intermediate care units as cardiac and non-cardiac may not have been a precise method to provide differentiation of units. Although this work supports that differences are found across these groups the patterns are not consistent or clear. Measurement of outcomes was conducted without reliability measurements because the outcomes were pulled from existing databases or surveillance methods that do not use these techniques. Some of the outcomes actually may have occurred on other units but when counted via prevalence methods were assigned to another unit. Under- and overreporting could have occurred for a number of reasons. Hospital or unit quality improvement efforts at enhancing error reporting, which occurred over the data collection period, might have allowed outcome rates such as medication errors to increase.
Research implications	Larger sample sizes are needed to investigate these relationships further. Does staffing impact on disease-specific outcomes (e.g. the rate of dysrhythmia detection)? Would the results still hold if adjustments were made? Do different units have the same experience, organisational and mix of staff? Need to incorporate interpersonal unit process variables (e.g. unit culture, communication, co-ordination, leadership etc.)

ID, origin, authors (year)	74, USA, Zimmerman, S., Gruber-Baldini, A.L., Hebel, J.R. *et al.* (2002)
Aims	To determine the relationship between a broad array of structure and process elements of nursing home care and resident infection and hospitalisation for infection. *Workforce:* Director of nursing (DON), Registered Nurses (RNs), license practical nurses (LPNs), aides, therapists, physicians, volunteers and administrators, nursing homes (NH) *Feature:* Staffing levels, experience and turnover *Outcome:* Infection (written diagnosis, a course of antibiotic therapy, or radiographic confirmation of pneumonia) and hospitalisation for infection (indicated on medical records)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, prospective cohort 2 Subjects were a new admission cohort of 2285 residents recruited from a sample of NHs participating in the Maryland Long-Term Care Project. 3283 new admissions were identified who had not previously resided in a long-term care facility for 8 or more days, were aged 65+, and for whom baseline data could be collected within 2 months of admission. Infections through the first 7 days after admission were excluded, as was onychomycosis of the toenail and infection indicated by prophylaxis orders for antibiotics if ordered 1 day before surgical procedure. 3 59 nursing homes, 2015 patients 4 2 years 5 Baseline data were collected from September 1992 through March 1995 and facility data were collected at the midpoint of follow-up. Facility-level data were collected from: interviews with facility administrators, directors of nursing, and activity directors; record abstraction; and direct observation. Resident-level baseline data were collected using chart abstraction and interviews with residents, care providers, and family members an average of 31 to 40 days after admission. Information on staffing was abstracted from the facility's Health Care Financing Administration Form 671. Data on the physical environment and resident and staff activity were conducted with the Therapeutic Environment Screening Survey for Nursing Homes (TESS-NH) and the Resident Staff Observation Checklist (RSOC).
Results Quantitative results	Only results that were significant predictors associated with the outcomes were reported. Administrator experience, DON experience, administrator turnover, DON turnover, RN FTE/100 beds/week, physician hours/100 beds/week, mental health hours/100 beds/week, volunteer hours/100 beds/week, LPN turnover/FTE, nursing aide turnover/FTE, and resident autonomy were all investigated but were not found to be significant. Risk factors for infection were: more therapist FTEs, more LPN FTEs and fewer nursing aides. Higher RN turnover was significant for both outcomes

Variable	**Infection**	**Hospitalisation for infection**
		Relative Risk (95% CI)
Therapist FTE/100 beds	1.03** (1.01–1.06)	1.06 (0.98–1.16)
Licensed practical nurse FTE/100 beds	1.85** (1.22–2.78)	3.08 (0.74–12.85)
Aide FTE/100 beds	0.86* (0.77–0.97)	0.82 (0.46–1.49)
Registered nurse turnover/FTE	1.29* (1.03–1.62)	1.83* (1.04–3.23)

*p <0.05, **p <0.01

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No 2 No 3 Yes 4 70% gave consent; follow-up data were available for 2015 residents; 80 were discharged before 8 days; 12 denied permission for follow-up; 178 had missing or incomplete data. 5 Stratified random sample 6 Statewide – Maryland
Commentary	To account for differences in resident characteristics between facilities, a model was fitted that included the facility variable and resident characteristics (i.e. sex, age, ethnicity, education, marital status, morbidities, functional status, and RUG III scores) as covariates. However, the authors state that in no case did the adjusted relative risk differ substantially from the unadjusted one, and thus only reported unadjusted figures. No account was made for the change of facility characteristics over time. It is likely that the incidence of infection is underreported due to the method of data collection chosen.
Research implications	Can patients be treated in a NH as opposed to being hospitalised? How can infection rates be reduced and which facility characteristics play a role in this? Staffing and resident acuity require further investigation in relation to infection rates, as do administration, policies and practices in regard to hospitalisation.

Table A2.3 Workforce to population ratio

ID, origin, authors (year)	684, UK, Gulliford, M.C. (2002)
Aims	To evaluate whether population health was associated with general practitioners supply in England *Workforce:* Doctors, primary care *Feature:* Availability *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, ecological 2 All causes of mortality at ages 15–64 years, infant mortality, 'avoidable' mortality from conditions amendable to medical intervention and for acute myocardial infarction 3 Not stated 4 Not stated 5 Data were obtained from the English Department of Health's statistical publications on population size, the number of whole-time equivalent GPs per 10,000 weighted population and the proportion of residents in households headed by persons born in the Commonwealth as a measure of the proportion of ethnic minorities.
Results Quantitative results	Infant mortality, all-cause mortality, avoidable mortality and mortality from AMI were all lower in areas with more GPs. After adjusting for deprivation score, social class and the proportion of ethnic minorities there was only weak evidence for an association between GP supply and mortality indicators. When limiting long-standing illness was included as an additional confounder, then the mortality indicators were not associated with GP supply.

Indicator	**Median (range)**	**Correlation with GP supply**	**Mean change (95% CI) per unit increase in GP supply (per 10,000)**			
			Model 1	**p-*value***	***Model 2***	**p-*value***
All-cause mortality 15–64 years (SMR)	89 (70, 154)	–0.68	–5.2 (–8.3, –2.0)	0.002	–3.3 (–6.7, 0.1)	0.060
Infant mortality rate (per 1000)	5.5 (2.7, 9.5)	–0.34	–0.4 (–0.9, 0.2)	0.154	–0.2 (–0.8, 0.4)	0.493
Avoidable mortality	98 (71, 148)	–0.55	–5.3 (–9.7, –0.8)	0.022	–4.2 (–9.2, 0.8)	0.095
AMI (SMR)	97 (39, 206)	–0.64	–10.3 (–19.3, –1.3)	0.026	–5.5 (–15.3, 4.3)	0.269

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 & 2 The Townsend score was used as a measure of deprivation. The score is based on the proportion of people in an area who are unemployed, living in overcrowded accommodation, not in owner occupation and not owning a car. The proportion of people in households headed by people in social class IV (semi-skilled manual occupation) and V (unskilled manual) was also included. Health indicators included the proportion of the population with limiting long-standing illness. Indirectly standardised hospital admission rates per 100,000 for acute conditions (infections of the ears, nose and throat, or kidneys and urinary tract and heart failure) and chronic conditions (diabetes and asthma), and conception rate per 1000 females <18 years were also included. Observations were weighted for health authority population size. 3 Yes 4 Not stated 5 No 6 99 health authorities in England

Commentary	The finding of higher mortality from all causes in areas less well supplied with primary care doctors is non-specific, and might perhaps result from confounding with wider influences on health. GPs choice of location may be very sensitive to the quality of environment and amenities in an area and the confounders included in this analysis are unlikely to fully account for the impact of deprivation on health. As well as having more GPs, more affluent areas have general practices with better facilities, providing more services and offering longer consultations with higher quality of care. This is an ecologic study that is subject to the ecologic fallacy, in which associations at the population level do not accurately reflect associations at the individual level. No information was available on individual patients' actual use of physician services. Ecologic studies have very limited ability to establish causation, and follow-up studies conducted at the individual patient level will be necessary to confirm these findings.
Research implications	Need and outcome cannot be distinguished in cross-sectional data and future longitudinal studies with improved adjustments are needed.

ID, origin, authors (year)	1126, USA, Goodman, D.C., Elliot, S., Fisher, M.D. *et al.* (2002)
Aims	To determine whether a greater supply of neonatologists or neonatal intensive care beds is associated with lower neonatal mortality *Workforce:* Neonatologists, neonatal intensive care *Feature:* Availability (very low, low, medium, high, and very high supply) *Outcome:* Neonatal mortality (death within the first 27 days of life)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, cohort 2 Infants with a birthweight of <500g were excluded because they are not always classified as live births. Of the 3199 physicians who reported themselves as neonatologists, those that spent the majority of their time teaching (97), doing administrative work (100) or research (232) and those working <20 hours per week (118) were excluded. 3 3,892,208 newborns, 246 neonatal intensive care regions, equivalent of 2407 full-time neonatologists (the total number of fellows was multiplied by 0.35 to adjust for less active clinical roles (0.35 x 377 = 132)) 4 27 days of life 5 Used linked birth and death records from the 1995 birth cohort to assess the associations between supply of neonatologists and neonatal intensive care beds per capita and the risk of mortality. Master files of the American Medical Association and the American Osteopathic Association and 1998 and 1999 surveys of neonatal intensive care units to calculate the supply of neonatologists and neonatal intensive care beds in different regions. The number of neonatal intensive care beds and intermediate care beds were determined using a survey.
Results Quantitative results	The numbers of neonatologists and beds were not consistently larger in areas where the need for neonatal intensive care was greatest. **Supply of neonatologists** / **Number of deaths per 1000 births** / **Adjusted odds ratio (95% CI)** Very low (2.7/10,000 births) / 3.5 / 1.00 (reference group) Low (4.3/10,000 births) / 3.3 / 0.93 (0.88–0.99) Medium (5.9/10,000 births) / 3.3 / 0.93 (0.88–0.99) High (7.5/10,000 births) / 3.4 / 0.91 (0.86–0.97) Very high (11.6/10,000 births) / 3.5 / 0.89 (0.83–0.95) The risk of neonatal death was lower in regions with a low supply of neonatologists than in regions with a very low supply. However, little additional benefit was seen with further increases in supply. Associations between a very low supply of neonatologists and an increased risk of death were limited to the infants with the lowest birthweight.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Controlled for birthweight, sex, type of birth (singleton or multiple), maternal age (<15, 15–19,20–29, 30–34 or 35>), parity (primiparous or multiparous), ethnicity (white, black or other), level of education (<12, 12, 12–15 or 16=>), marital status (married or unmarried), and extent of prenatal care (none, beginning in first trimester, beginning after the first trimester, or unknown) and clustering of neonatal mortality within regions 2 None 3 Source of the data collection was uniform but the period was not. 4 Not stated 5 No 6 Not stated

Commentary	Lots of the data were not reported and there was no reference to where this information could be found. The data on the number of neonatologists and neonatal intensive care beds were not from the same years as the birth cohort studied. Data on the health status in infancy other than mortality or long-term outcomes were not studied. Intensive care resources are measured at the regional level and there were no data on the process of care; the actual causes of higher mortality rates in regions with the lowest resources cannot be determined. The authors neglect to consider the real-world implications of fewer neonatologists. They also fail to take into account the distribution of work time by clinically active neonatologists, but believe that there is no evidence that this additional work varies significantly in relation to the variation in the supply of neonatologists.
Research implications	Would infants benefit from the greater availability of neonatologists and resources in ways that are not reflected by mortality (e.g. lower morbidity)? Further research is needed to identify meaningful measures of outcomes other than mortality that may be sensitive to differences in the regional supply of specialists, as well as to identify possible reasons why increases in supply may not produce improvements in health. Would infants in regions with more neonatologists receive more attentive care, resulting in faster resolution of illness, lower rates of complications, and better subsequent health status than infants in regions with fewer neonatologists? Alternatively, in regions with a greater supply of neonatologists, would infants with less serious illness be more likely to be admitted to a neonatal ICU and subjected to more intensive diagnostic and therapeutic measures, with the attendant risks of errors and iatrogenic complications, as well as impaired family–infant bonding? Do the following affect the results: the volume of very sick infants cared for in neonatal ICUs; the level of care provided to high-risk newborns; the experience and training of the workforce; the skill mix and grade mix of the workforce; the specific treatment provided; and delays in initiating care because of the need to transfer neonates rather than treat them in the hospital they were born?

ID, origin, authors (year)	1128, UK, Jarman, B., Gault, S., Alves, B. *et al.* (1999)
Aims	To ascertain hospital inpatient mortality in England and to determine which factors best explain variation in standardised hospital death ratios *Workforce:* GPs, nurses, mixed settings *Feature:* Ratios – hospital doctors to beds, GPs to head of population *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, retrospective 2 *Inclusion:* Discharge records only (episodes that ended in discharge (alive or dead) from the hospital rather than transfer to the care of another consultant within the hospital) for primary diagnosis of one of 85 primary diagnoses which accounted for 80% of deaths *Exclusion:* Community and speciality institutions, small hospitals (under 9000 admissions during the 4 years) and hospitals without accident and emergency units 3 183 acute general hospital trusts 4 In-hospital 5 Three main sources: the NHS hospital episode statistics data system from 1991/2 to 1994/5, the national decennial census from 1991 and other routine NHS data such as hospital characteristics, hospital staffing levels and general practitioner distribution.
Results Quantitative results	Weighted multiple linear regression using two models: A included all admissions both elective and emergency; B included emergency admissions only. Stratified by age (using 10-year age groups), sex and the 85 primary diagnoses. Aggregate discharge data were taken from individual records and aggregated across each hospital. Community data were taken from geographical areas attributed from area of residence to each discharge (via postcode) and then averaged across discharges for each hospital. **Model A** After adjustments for the % of emergency admissions, the best predictors of hospital mortality were numbers of hospital doctors per 100 hospital beds and general practitioners per 100,000 population. Higher hospital standardised mortality ratios were associated with lower numbers of hospital doctors per hospital bed and lower numbers of GPs per head of the population. A reduction in 5000 hospital deaths per year was associated with a 27% increase in hospital doctors or an 8.7% increase in general practitioners. **Variable** — **Regression coefficient (95% CI)** — ***p*-value** — **Mean** Number of hospital doctors per 100 hospital beds in 1994/5 — –0.47 (–0.64 to –0.30) — <0.001 — 25.4 (8.0) Number of general practitioners per 100,000 population — –0.67 (–1.05 to –0.30) — <0.001 — 54.6 (3.4) **Model B** At the 5% level of significance, the proportion of grade A nurses (auxiliary nurses in training) as a percentage of all hospital nurses and bed occupancy entered the model. High percentages of grade A nurses were associated with higher hospital standardised mortality ratios. **Variable** — **Regression coefficient (95% CI)** — ***p*-value** — **Mean** Number of hospital doctors per 100 hospital beds in 1994/5 — –0.51 (-0.65 to –0.38) — <0.001 — 25.4 (8.0)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 *Included in multiple regression analysis:* Aggregate discharge data: percentage of emergency cases; percentage of cases and deaths with co-morbidity of the 85 diagnoses leading to 80% of all deaths and combinations of those with the highest correlations with hospital standardised mortality ratios; percentage of cases and deaths with each of the top 15 diagnoses which account for 50% of deaths; percentage of cases with co-morbidity of the two or three conditions most highly correlated with hospital mortality; health authority where hospital located. Hospital data: hospital doctors per bed; percentage of nurses at grade A; bed occupancy; location (inner London and outer London); university teaching. Community attributed data: general practitioners per 100,000 population according to ONS based on health authority of patient residence; NHS facilities per 100,000 population in hospital local health authority; underprivileged area score; one-parent families; mobility. *Other independent variables included in univariate analysis:* Aggregate discharge data: percentage of live discharges who went home; average number of diseased bodily systems; average LOS; number of cases. Hospital data: hospital doctors per case; percentage of nurses above an A grade; nurses per doctor and per bed; number of hospital beds; percentage of geriatric beds; location outside of London; non-university teaching; other general hospital; provision of a range of specialist units; hospital income per bed and per case; total and first accident and emergency attendances; hospital character standards; results of survey of patient-centred care. Community attributed data: general practitioners per 100,000 population according to ONS based on individual data averaged at health authority of residence level; general practice nurses per 1000 population according to ONS in hospital local health authority; elderly living alone; children aged under 5; social class V; unemployed; overcrowded accommodation; ethnic minority; percentage of patients with limiting long-standing illness; provision of nursing homes and residential care homes in hospital local health authority area 3 Three sources 4 Excluded hospitals that had poor-quality data, more than 30% of inpatient episodes without a valid discharge or more than 30% of primary diagnoses recorded as unknown 5 No – used criteria based on type and size as well as quality of data recorded in the Hospital Episode Statistics (HES) database 6 Two hospitals per health authority across England, 85% of all admissions in England HES database
Commentary	Only one of the measures of co-morbidity was significant in the model and this might be related to the lack of data on the severity of illness. The figures used were the aggregates for the health authority of hospital location rather than individual figures for each hospital's emergency catchment area (often very different). It could simply be that more doctors mean more admissions. The data are an inadequate basis for drawing the overall conclusion that higher hospital doctor and GP ratios equates to lower death rates. The accuracy of the measurement of prognostic or risk factors in patients treated by the hospitals is inadequate. The rates derived from the HES database represent episodes rather than actual patients. Hospital doctors and acute care beds per 100,000 population would be better examined independently. There are other factors outside of the hospital that could affect hospital death rates after the adjustments made in this instance, such as care in the community.
Research implications	Repeat the above analysis with data aggregated by electoral ward of residence rather than by hospital of admission. What about nurses, do they have the same impact on mortality?

ID, origin, authors (year)	505, UK, Robinson, J. and Wharrad, H. (2000)
Aims	To explore the relationship between the global distribution of health professionals, gross national product per capita, female literacy and the health outcome indicators of infant, and under-5 mortality rates using available data from United Nations' sources *Workforce:* Nurses and physicians *Feature:* Ratios of health services personnel to population *Outcome:* Infant mortality (IM, the number of deaths of infants under 1 year of age per 1000 live births) and under-5 mortality rates (u5MR, the probability that a newborn will survive to exactly age 5, based on prevailing mortality rates)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, ecological 2 All countries with the required data available 3 The number of countries varies between 109 and 155 in the analysis 4 Unsure 5 United Nations (UN) sources were used to provide data for physicians (1990–1995) and nurses (1988–1992) per 1000 population. The United Nations Children's Fund (UNICF, 1997) was used to find the % of female literacy (1995). The IMR and u5MR data for 1995 were collected from the World Bank (1997).
Results Quantitative results	*Infant mortality rates:* Based on data from 148 countries for physicians and 149 for nurses. The correlation coefficients are –0.80 for physicians and –0.72 for nurses (p <0.001) showing that countries that have high IMRs also have low numbers of physicians and nurses. Multiple regression analysis shows that 2% of the variation in IMR is associated with physicians per 1000 population. Nurses added nothing. *Under-5 mortality rates:* Based on data from 148 countries for physicians and 149 for nurses. The correlation coefficients for u5MR are –0.81 for physicians and –0.72 for nurses (p < 0.001) showing that countries having high u5MR have low numbers of physicians and nurses. Multiple regression analysis shows that physicians per 1000 population are associated with 66% of the variation in u5MR. Nurses per 1000 population add nothing further to the outcome.

Dependent variable	**Independent variable**	**R-squared**	**Constant**	**Coefficient**	**ANOVA (p-value)**
Infant mortality rate	GNP	67% (22)	3.259 (0.117)	–0.559 (0.038)	< 0.00
Infant mortality rate	GNP	80% (18)	2.802 (0.107)	–0.367 (0.038)	< 0.00
	Female Literacy			–0.260 (0.032)	
Infant mortality rate	GNP	81% (17)	2.575 (0.126)	–0.318 (0.040)	< 0.00
	Female Literacy			–0.180 (0.040)	
	Physicians/1000			–0.129 (0.041)	
Under–5 mortality rate	Physicians/1000	66% (25)	1.559 (0.027)	–0.497 (0.034)	< 0.00
Under–5 mortality rate	Physicians/1000	79% (20)	2.178 (0.146)	–0.297 (0.037)	< 0.00
	GNP			–0.366 (0.046)	
Under–5 mortality rate	Physicians/1000	83% (18)	2.744 (0.132)	–0.157 (0.043)	< 0.00
	GNP			–0.327 (0.042)	
	Female Literacy			–0.209 (0.042)	

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Gross national product per capita (GNP) and % of female literacy were included in the analysis to take into account socioeconomic variables. 2 None 3 Yes 4 Unclear 5 No 6 155 countries
Commentary	This is an ecologic study that is subject to the ecologic fallacy, in which associations at the population level do not accurately reflect associations at the individual level. The exploratory nature of selecting variables for ecologic studies may also increase type 1 statistical errors, falsely concluding that associations exist when they have actually occurred by chance. In an earlier paper the authors found that 70% of the global variation in the distribution of nurses was associated with the distribution of physicians. These findings suggest that nurses are not normally substituted for more expensive physicians. Instead, countries elect first to employ physicians as the primary health care providers and then employ nurses in relation to physician numbers. Standardised residuals for the multiple linear regressions of IMR and u5MR, with GNP plotted on a scatterdiagram against the standardised residuals for the multiple linear regression of ratios of nurses with GNP and for ratios of physicians with GNP are included in the report. All the UN sources used for data identification emphasised the variable reliability of the data sets provided. Physicians and nurses per 1000 population were selected in this paper as surrogates for health services input. There is no way of telling from the data whether the personnel are deployed in acute hospitals or primary care, public or private sector, community general practice or high-technology specialities. The definition of a nurse is also problematic. Not only does it fail to distinguish between nurses and midwives, but also there is no global standard definition of a registered nurse. The 'disappearance' of nurses as an associated variable from each of the multiple regression analyses raises issues of validity. As it is nurses who administer immunisations and, in many developing countries, represent the only qualified personnel in community health centres, this finding appears to be an anomaly. The use of GNP per capita as a measure of a country's wealth has several limitations. It provides no indication of the degree of equity with which wealth is distributed across a country's population. Equality of income is better expressed by the Gini coefficient but this indicator was unavailable for many countries. Further, GNP does not distinguish between the aims and ultimate uses of a given product or harms or contributes to welfare. GNP also varies by climate and does not deal adequately with environmental issues. The data for physicians and nurses were collected over different periods.
Research implications	Individual countries within this study should conduct case studies of their respective situations to establish the reliability of the findings.

ID, origin, authors (year)	653, USA, Roetzheim, R.G., Gonzalez, E.C., Ramirez, A. *et al.* (2001)
Aims	To determine if increasing primary care physician supply was associated with lower incidence and mortality rates of colorectal cancer. *Workforce:* Physicians, primary care (classified as primary care if their self-designated speciality was family practice, general practice, obstetrics/ gynaecology or general internal medicine) *Feature:* Supply (total physician supply, primary care physician supply and non-primary care physician supply) *Outcome:* Incidence and mortality rates for colorectal cancer (stratified by proximal cancers and distal origin of the cancer)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, ecological 2 Physicians who indicated they were either 'semi-retired', in residency training or engaged in teaching or research counted as 0.5 full-time equivalent (FTE). Physicians who indicated they were no longer involved in direct patient care were excluded. 3 Not stated 4 In-hospital 5 Incidence and mortality rates were identified using the Florida Cancer Data System (FCDS). 1990 US census data were used to ascertain other characteristics of Florida counties that might have an impact on colorectal cancer incidence and mortality. Data on physician supply were obtained from 1994 American Medical Association physician master file. County-level population estimates were obtained from the 1990 US Census. Data collection: 1993–1995.
Results Quantitative results	Increasing the supply of primary care physicians was associated with lower incidence and lower mortality rates of colorectal cancer in Florida counties. Each 1% increase in primary care physician supply was associated with a reduction in colorectal cancer incidence of 0.25 cases per 1,000,000 and a reduction in mortality of 0.08 cases per 100,000. In contrast, overall physician supply was unrelated to the outcomes. **Rates** — **Correlation coefficients** — ***p*** Total incidence rates — –0.46 — <0.0001 Proximal cancers — –0.48 — <0.0001 Distal cancers — –0.36 — 0.003 Total mortality rates — –0.29 — 0.02 *Results of linear regression analyses* **Parameter estimate (SD)** — ***p*** — **95% CI** Total incidence rates — –24.8 (5.8) — <0.0001 — –36.5 to –13.1 Proximal cancers — –11.6 (2.7) — <0.0001 — –17.0 to –6.2 Distal cancers — –13.4 (4.3) — 0.003 — –22.0 to –4.8 Total mortality rates — –7.8 (2.2) — 0.0008 — –12.2 to –3.4
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 The FCDS provides age-adjusted incidence and mortality rates by standardising them to the 1970 US standard population. To account for year-to-year fluctuations, rates were averaged over the 3-year period 1993 to 1995. Variables obtained from each county included median household income, percentage of county residents with less than high school education, percentage residing in urban census areas, percentage who were white, and percentage who were married. 2 None 3 Yes 4 Not stated 5 No 6 67 Counties in Florida

Commentary	This is an ecologic study that is subject to the ecologic fallacy, in which associations at the population level do not accurately reflect associations at the individual level. No information was available on individual patients' actual use of physician services. Ecologic studies have very limited ability to establish causation, and follow-up studies conducted at the individual patient level will be necessary to confirm these findings. The exploratory nature of selecting variables for ecologic studies may also increase type 1 statistical errors, falsely concluding that associations exist when they have actually occurred by chance. Other risk factors for colorectal cancer (such as dietary patterns, family history or rates of ulcerative colitis) were not considered. As the outcomes were established according to the patient's county of residence rather than location of diagnosis or treatment the authors don't believe the associations observed were the result of referral patterns. However, physician supply might be correlated with other unmeasured characteristics of the health care delivery system, which could account for the observed associations. This study was restricted to colorectal cancer in Florida, which may not be representative of other diseases or countries.
Research implications	Studies conducted at the individual patient level are necessary to confirm the findings.

Table A2.4 Workforce hours

ID, origin, authors (year)	392, USA, Bliesmer, M.M., Kane, R.L. and Shannon, I. (1998)
Aims	To explore the relationship between measures of resident outcomes and nursing home characteristics: size, ownership, licensed and non-licensed nursing hours, and compliance with a state correction order *Workforce:* Licensed nurses (Registered Nurses (RNs), licensed practical nurses (LPNs)) and non-licensed nurses (e.g. nursing aides), nursing homes *Feature:* Workforce hours *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, historical cohort study 2 Nursing home residents aged 65 and older admitted in the study period 3 4103 patients for 1988; 4676 for 1989; 4672 for 1990; 440 nursing homes 4 3 years in institution 5 The data on patients' health outcomes and nursing hours' were from the records during the study; data on nursing homes and nursing hours per standardised resident day were from the Minnesota Department of Human Services Long-Term Care Division facility profile from 1988 to 1990.
Results W—F**?** O Quantitative results	More nursing hours were associated with a lower risk of death. **1989** **1990** **1991** Licensed nursing hours –0.00079*** –0.00091*** –0.00105*** Non-licensed nursing hours 0.00016 –0.00014* –0.00024** *p <0.05, **p <0.01 and ***p <0.001
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Adjusted for patients' age, gender, admission Total Dependence Score (TDS) 2 Adjusted for the nursing home size, ownership, and the compliance with a state correction order 3 Uniform data based on Minnesota case mix reimbursement system 4 Not stated 5 50% selection of available patients of 1988 admissions, 20% for both 1989 and 1990 admissions 6 One state
Commentary	Some of element of case mix was not captured in the measures of age and functional status on admission, i.e. behavioural problems. Only Minnesota nursing home resident and facility data were analysed. The study used an administrative data set to collect the reimbursement, regulation, and cost factoring of Minnesota nursing home care. Because of the nature of these data, it could not separate the effect of benefits from more active professional nursing that occurs immediately after admission from those that occur later in the patient' course. Outcomes were accessed only annually. It is possible that most of the benefit is enjoyed by those receiving what amount to subacute care. All variables that might have influenced the outcomes of nursing home residents were not known and therefore could not be measured or controlled in this study.
Research implications	Need a new paradigm for quality rather than the traditional techniques that tend to equate good nursing home care simply with the absence of untoward events. An outcome-oriented approach may provide an environment for communication and consultation between regulatory agencies and nursing homes concerning the actual intention of public policy. Further studies are needed that focus on the relationship between staffing levels and patient outcomes while controlling for other factors.

ID, origin, authors (year)	660, USA, Laine, C., Goldman, L., Soukup, J.R. *et al.* (1993)
Aims	To examine the impact on patient care of a New York State regulation that restricted house staff working hours. Code 405 went into effect on July 1 1989 and dictated that house officers could work a maximum of 80 hours per week, with a maximum of 24 consecutive hours of patient care and a minimum of 8 hours off duty between shifts; 3 hours of overlap to exchange information were permitted after a 24-hour shift but during this period direct patient care was forbidden. *Workforce:* Medical house staff, teaching hospital *Feature:* Workforce hours; restriction of working hours *Outcome:* Mortality, LOS and in-hospital medical complications
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, two retrospective cohorts – one cohort discharged during October 1988 before Code 405 and one cohort discharged from the same service during October 1989 after Code 405 2 Control: 18 patients were excluded due to incomplete data; post-Code 405: 20 patients excluded due to incomplete medical records 3 18 house officers staffing the service before and after the implementation of the Code 4 In-hospital 5 The medical records for each patient hospitalised were reviewed.
Results W—F**?** O Quantitative results	A univariate comparison of the outcomes showed no differences between the two cohorts in in-hospital mortality, mean LOS and most of the complications.

Outcome	**1988 cohort**	**1989 cohort**	***p*-value**	**After adjustment**
Mortality	25 (9.5)	25 (9.5)	RR = 1.0 (0.59–1.7)	
LOS	9.55	9.49	0.94	
Toxic/metabolic drug reaction	10	14	not significant	
Anaphylactic drug reaction	0	2	not significant	
Drug-related rash	6	8	not significant	
Transfusion reaction	2	1	not significant	
Urinary tract infection	10	15	not significant	
Pneumonia	7	9	not significant	
MRSA Infection	1	5	not significant	
Wound Infection	3	1	not significant	
Line Infection	2	3	not significant	
Other in-hospital infection	7	11	not significant	
Pulmonary embolism	0	3	not significant	
Deep vein thrombosis	0	2	not significant	
Renal insufficiency	11	14	not significant	
Electrolyte abnormality	8	21	0.01	
Respiratory decompensation	8	11	not significant	
Other in-hospital complication	17	33	0.02	
Total complications	92	153	0.001	
Patients with at least 1 complication	59	91	0.002	OR = 1.9 (1.2–3.0)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Age; sex; number of co-morbidities and medications on hospital admissions; Charlson co-morbidity score and stratum; whether the patient was admitted from the emergency department, an office, another institution, or home; recent prior hospitalisation; whether the patient was admitted via an overnight holding service; whether the patient was located on a floor other than the home floor of the covering house staff team; and, for test delays only, the number of tests ordered. 3 Yes 4 Not stated 5 No 6 One general internal medicine teaching service of The New York Hospital
Commentary	Although the New York State enacted Code 405 to prevent adverse patient outcomes, the authors found no such benefit. While the most serious outcomes such as in-hospital mortality were statistically unchanged, in-hospital complications were more frequent after the regulation. This could be a consequence of the power of this study as it was too low to detect differences in less frequent more serious outcomes. For example, there was only a 41% power of detecting a 5% change in mortality. If the authors had waited longer before studying the impact of Code 405 the hospital might have had more time to adjust, and perhaps the inefficiencies that were found would no longer be present. Because the supervision on the study service was already at the level mandated by Code 405 before the regulation went into effect, the authors could only look at the working-hour portion of the regulation. A study of a hospital or a service where supervision had been below requirements before Code 405 might show improvement in quality following the regulation. The authors note that the small number of house staff studied and the variation in the quality of the house staff between the study periods could account for the findings. The findings of this study may not be generalisable outside the study population as the staffing and educational requirements of different specialities at different institutions vary substantially. Another issue that limits the generalisability of this study is that there are numerous ways to comply with the regulations. Hiring more house staff, creating patient units covered by attending physicians or fellows rather than by residents, supplying greater clerical and ancillary support, adopting shifts similar to those of nursing staff, instituting night float teams, or a combination of these strategies could have fulfilled the requirements.
Research implications	Does the patient-to-physician ratio contribute more to the quality of care than does continuity of care or well-rested physicians? Which strategies (as above) optimise the quality of patient care? Does the skill mix of the staff on the wards make a difference?

Table A2.5 Maternity outcomes

ID, origin, authors (year)	680, Australia, Callaghan, L.A. *et al.* (2003)
Aims	To investigate the association between infant to staff ratios and the outcome of very low-birthweight (VLBW) babies *Workforce:* Secondary care setting; Nursing workforce: neonatal intensive care nurses *Feature*: Staffing levels of nurses related to health outcomes *Intervention/comparison:* The study examined the number of infants per nurse per shift on a neonatal intensive care unit (NICU) and related it to the risk of mortality using regression analysis. Infant to staff ratios resulting in the highest and lowest mortality levels were determined and significant differences found that were not explained by unit dependency or case mix. *Outcomes:* Infant – primary outcome measure was mortality before discharge.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective observational study 2 Infants with a lethal congenital abnormality were excluded along with infants not admitted to the ICN within 36 hours of birth. 3 709 VLBW admissions during study period. 17 excluded from analysis due to lethal congenital malformations (n=4), and delayed admission to NICU (n=13). Final size of study population n=692. 4 Admission to NICU within 36 hours postpartum until discharge or death 5 Data of infant numbers, infant characteristics of the birth history, admission details and physiological data of the first 12 hours of life were extracted from routine records at the hospital site. The period of interest was from January 1996 to December 1999. Data of the number of nurses working every shift were collected from routine records on a separate database for the same period. Dependency data, categorising individual patients as intensive care, high or medium dependency or recovery were also collected. Data not available from the database were retrieved by chart review.
Results Quantitative results	Outcome by infant:staff ratios, pooled over 9 shifts (divided into terciles) with crude odds ratios (OR) of mortality and adjusted mortality OR (adjusted for Clinical Risk Index for Babies (CRIB) only, CRIB and dependency) **Infant:staff pooled ratios** / **CRIB mean (SD)** / **Survived** / **Died (% total)** / **Total** / **Crude OR (95% CI)** / **Adj OR (95% CI) (CRIB)** / **Adj OR (95% CI) (CRIB and dependency)** 1.16–1.58 / 3.79 (4.05) / 202 / 30 (12.9) / 232 1.59–1.70 / 4.40 (4.27) / 193 / 37 (16.1) / 230 / 1.29 (0.77 to 2.17) / 1.05 (0.55 to 2.00) / 0.84 (0.42 to 1.66) 1.71–7.97 / 3.71 (3.69) / 217 / 13 (5.7) / 230 / 0.40 (0.21 to 0.80) / 0.32 (0.15 to 0.71) / 0.18 (0.06 to 0.5) Overall hospital mortality rate was 12% (80 out of 692 infants) *Summary:* An increase in the infant:staff ratio does not increase the risk of mortality. In comparison with the lowest infant:staff ratio, the odds of mortality are reduced.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 CRIB was used to adjust for case-mix. CRIB is used to provide a score for initial neonatal risk in infants <1500g, or <31weeks gestation. Of the population who died (n=80), the mean CRIB score was 9.98. Of the population who survived (n=692), the mean CRIB score was 3.18 (p <0.001) 2 Adjustment was also made for dependency, calculated using the expected number of nurses for each shift. 3 No gaps in data collection reported. 4 Not applicable 5 Not applicable 6 One neonatal intensive care unit in a Brisbane maternity hospital
Commentary	Clinical characteristics of the study population were presented. Place of birth of the baby was discussed as a possible confounder; however, this was not found to be significant.
Research implications	The authors state this study suggests the need for further investigation into infant:staff ratios as an independent risk factor in the mortality of VLBW babies and suggest that a prospective study, examining both total infant and staff numbers plus individual nursing workloads and perceived staffing requirements versus actual staffing supplied could clarify the impact of staffing on infant mortality.

ID, origin, authors (year)	199, UK, Robinson J.A. and Wharrad, H. (2001)
Aims	To consider the relationships between the global distribution of health professionals – ratios of physicians and nurses to population, and the proportion of births attended by trained health professionals – and gross national product (GNP) per capita, female literacy and maternal mortality rates (MMR) *Workforce:* Mixed primary and secondary care setting; mixed workforce: doctor, nurse, midwife *Feature:* Staffing levels in terms of the distribution of health professionals *Intervention/comparison:* A study examining the relationship between density of doctors, nurses and trained birth attendants and the GNP, female literacy, and MMR. Multiple regression analyses were performed using number of health professionals, GNP and female literacy as explanatory variables and MMRs as the dependent variable. *Outcomes:* Maternal – distribution of physicians and nurses and attendance at births; MMR
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective observational study 2 None stated 3 Total number of countries entered into regressions varied between 112 and 144. 4 Data covered varying periods (see below) 5 Data for health professionals came from the UN, and further data were added to this to make a combined database. (a) Physicians (1990–1995) and nurses (1988–1992) per 1000 population and GNP per capita (US$, 1995) from UN data sources (b) Percentage female literacy (1995) from UNICEF (1997, 116 countries) (c) Maternal mortality rate (1990) from Revised Estimates of Maternal Mortality (WHO and UNICEF, 1996), total 145 countries (d) Percentage of births attended by physicians, nurses, midwives or primary health care workers trained in midwifery (1990–1996) from UNICEF (1997, 119 countries)
Results Quantitative results	Maternal outcomes: distribution of physicians and nurses and attendance at births Summary of linear regression analysis for GNP, physicians per 1000 population, nurses per 1000 population, female literacy and births attended against MMRs:

Dependent variable	**Explanatory variable**	**No. of countries**	**Adj R^2**	**Constant**	**Coefficient**	**ANOVA**
GNP	MMR	143	0.70	4.956	-0.887	F = 340, p <0.001
Female literacy	MMR	115	0.62	2.625	-0.757	F = 190, p <0.001
Physicians/1000	MMR	136	0.73	1.918	-0.921	F = 360, p <0.001
Nurses/1000	MMR	137	0.56	2.304	-0.870	F = 176, p <0.001
Births attended	MMR	118	0.83	3.395	-1.371	F = 570, p <0.001

The size of the positive correlation coefficient for MMRs (Adj R^2) against the number of physicians per 1000 population is 0.73 and 0.56 for MMR against number of nurses per 1000 population. These significant results show that countries having high MMRs also have low numbers of physicians and nurses.
Stepwise multiple linear regression analyses for GNP, Births attended and Physicians per 1000 pop`n on MMRs, Nurses per 1000 pop`n and, Female literacy:

Dependent variable	Explanatory variable	Adj R^2	Constant	Coefficient	ANOVA
MMR	Births attended	79%	3.343	-1.231	F=330, P < 0.001
MMR	Births attended GNP	85%	4.252	-0.845 -0.406	F=256, P < 0.001
MMR	Births attended GNP Physicians/1000	87%	6.677	-0.660 -0.297 -0.241	F=203, P < 0.001

The percentage births attended is associated with 79% of the variation in MMR (Adj R^2), a further 6% is associated with GNP per capita, and for physicians/1000 population a further 2%.

Maternal mortality rate

Figure 1 Scatterplot of physicans/1000 population

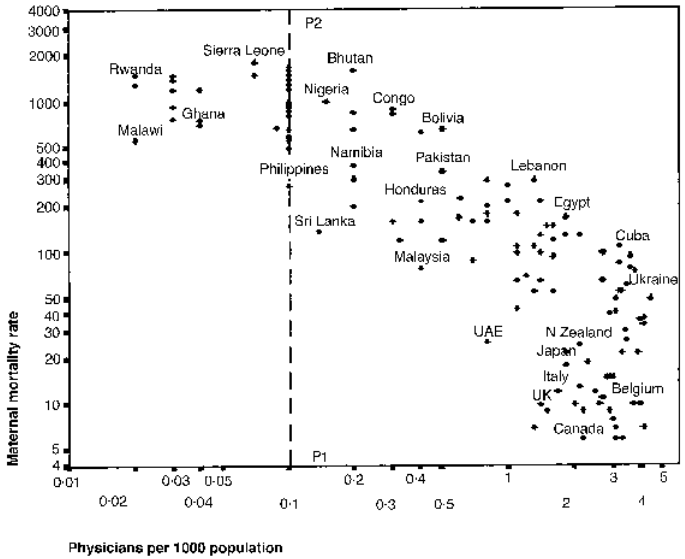


Figure 2 Scatterplot of nurses/1000 population

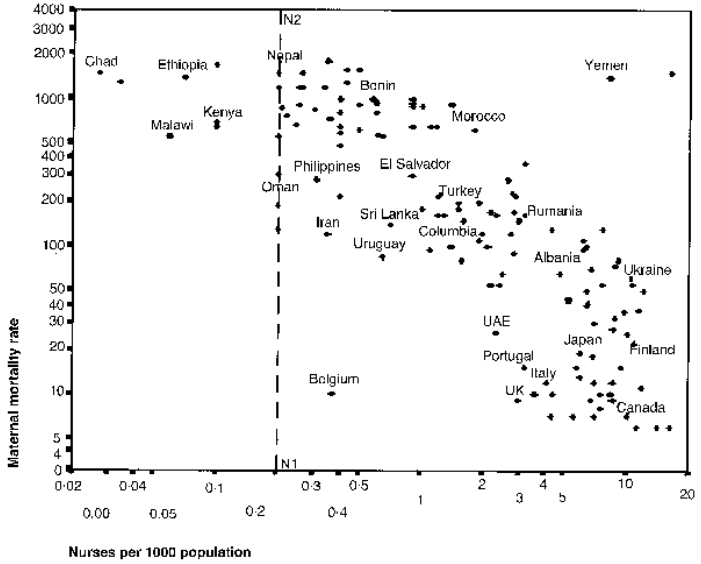
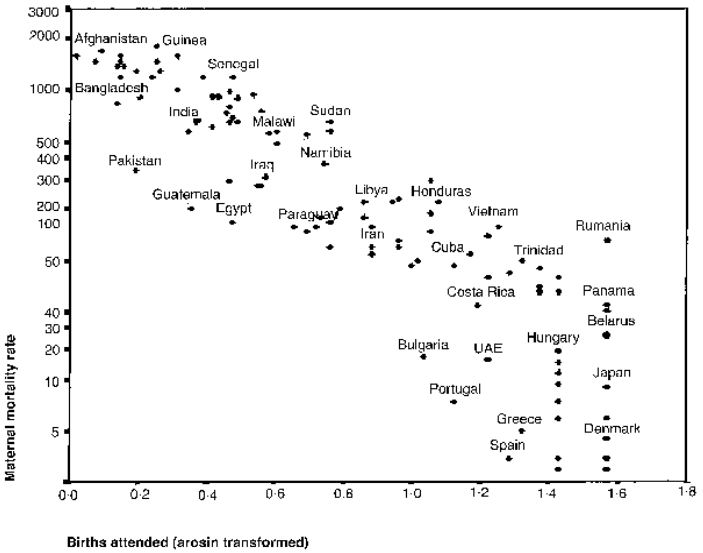


	Figure 3 Scatterplot of births attended by trained personnel and MMRs  *Summary:* Strong positive associations have been shown to exist between the ratios of physicians and nurses to the population, and the proportion of trained health personnel in attendance at birth, and MMRs. High numbers of health personnel correlate with low maternal mortality and vice versa
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Statistical adjustment was made for the explanatory variables – health professionals, GNP and female literacy. 2 Nonlinearity of the relationship between the variables was adjusted for by log transformation of the data for explanatory variables GNP, physicians per 1000 population and nurses per 1000 population. 'Attendance at birth' and female literacy percentage data was transformed to adjust for differences in underlying variances. 3 Gaps in data collection reported for: (a) UN data sources for GNP – GDP (US$, 1994) used for missing variables (b) Revised estimates of maternal mortality for maternal mortality rate – World Bank (1997, Table A6) used for missing variables. 4 None, records not linked 5 None 6 155 countries total – variable numbers of countries have been used in the regression analyses.

Commentary	GNP and literacy were used as socioeconomic indicators. The reliability of the UN data is variable and caution is asked for in interpreting the results – it is difficult to assess levels of maternal mortality at the national level as the data source for the MMR states, since reproductive age, cause of death and pregnancy status at time of death are all required information which is not reliably captured in many countries' data forms. The authors discuss why the variable nurses per 1000/population adds nothing further to the outcomes.
Research implications	The authors suggest that case studies of individual countries' respective situations to establish the reliability of these findings and to determine where to target the worst areas with highest MMRs. Data collection at the national level must be improved for many countries and health relief agencies must be proactive in this step to improve the quality of future research on the contribution of health professionals to reducing MMRs.

ID, origin, authors (year)	321, UK, Stilwell, J. *et al*. 1988
Aims	To investigate whether the mortality risk of babies is related to the resources available for their care at the time of birth To identify sources of routinely collected statistics of use when perinatal services at different units or districts are being monitored and compared *Workforce:* Secondary care setting; nursing workforce: nurse; midwife; medical workforce: obstetrician; paediatrician *Feature:* Staffing levels of medical and nursing staff *Intervention/comparison:* To compare the variation in the health outcomes of low-birthweight infants with the variation in staffing levels of midwives and nurses (MN), midwives (SCM) obstetricians (OB) and paediatricians (PD) of regional neonatal units over time. Regression analyses were performed using staffing measures and birthweight distribution as the explanatory variables and measures of mortality as the dependent variables. Outcomes, infant: stillbirths; neonatal deaths ≤6 days after birth Relevant results abstracted for the selected explanatory variable PD/LBW – Paediatricians proportionate to low Birthweight – for available years only
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective observational study 2 Inclusions were: stillbirths; early neonatal deaths. Exclusions were: babies over 2500g and deaths from malformations. Multiple births were excluded from the final analysis. Isolated GP units and regional neonatal referral centres were also excluded plus one maternity unit that did not contribute to the Hospital Activity Analysis (HAA) system. 3 20 maternity units. Total number of patients surveyed in each year unknown. 4 Six years activity reviewed. 5 Data collected for years 1977 to 1983. Staff measured by averaging the numbers of whole-time equivalent staff in post on 30 March and 30 September in each year of the study. Questionnaire completed by senior midwife on the unit's staffing in that period, about facilities, access to service within and beyond their own hospital site. HAA data were used to gather information on: birthweight; presence of malformation; birth outcomes. For two hospitals which did not contribute to the HAA data system, data were collected directly from the hospital. Each maternity unit supplied information about staffing, although records of medical staff by sub-specialisation – neonataologist or other paediatrician were not distinguishable. Also surveyed were resources and services: availability of 'on-site' and 'off-site' services, including 'call out' and travel times, and a list of items of equipment available in obstetric and neonatal units.

Results
Quantitative results

Key

PD/LBW	Paediatricians *proportionate to* low birthweights
PNMR	Perinatal mortality rate
PNMRCA	Perinatal mortality rate *not* congenital abnormality
FWD	First week deaths
FWDCA	First week death *not* congenital abnormality

Infant outcomes

Factors affecting infant outcomes

Relevant results abstracted for the selected explanatory variable PD/LBW for available years only.

Tables show the strength R^2 , and the significance *P*, of the correlation for each dependent variable.

Table 1 Variable PNMR: stillbirths and deaths ≤6 days rate recorded at hospital of birth per 1000 total births at that unit

Year	No of units	Variable X_1	Variable $b_1 \pm s_1$	*P* (*F*)	R^2	Adj R^2
1978	15	PD/LBW	–301.63 ± 88.58	0.005	0.471	0.431
1982	16	PD/LBW	–205.53 ± 67.94	0.009	0.395	0.352

Table 2 Variable PNMRCA: stillbirths and deaths ≤6 days rate for non-malformed singleton and multiple births per 1000 total births at that unit

Year	No of units	Variable X_1	Variable $b_1 \pm s_1$	*P* (*F*)	R^2	Adj R^2
1978	15	PD/LBW	–268 ± 82	0.006	0.45	0.41
1982	16	PD/LBW	–195 ± 61	0.006	0.42	0.38

Table 3 Variable FWD: deaths ≤6 days rate per 1000 total births

Year	No of units	Variable X_1	Variable $b_1 \pm s_1$	*P* (*F*)	R^2	Adj R^2
1977	16	PD/LBW	–198.80 ± 41.52	0.0003	0.620	0.594
1978	15	PD/LBW	–206.00 ± 53.33	0.002	0.534	0.499
1980	18	PD/LBW	–191.30 ± 57.21	0.004	0.411	0.375
1982	16	PD/LBW	–177.48 ± 39.21	0.0005	0.594	0.565

Table 4 Variable FWDCA: deaths ≤6 days rate for non-malformed singleton and multiple births per 1000 livebirths at that unit

Year	No of units	Variable X_1	Variable $b_1 \pm s_1$	*P* (*F*)	R^2	Adj R^2
1977	15	PD/LBW	–196 ± 46	0.0009	0.58	0.55
1978	15	PD/LBW	–189 ± 53	0.0034	0.49	0.46
1980	18	PD/LBW	–157 ± 50	0.0068	0.38	0.34
1982	16	PD/LBW	–170 ± 36	0.003	0.62	0.59

Table 5 West Midlands maternity units grouped by paediatric staffing ratios and 'in-house' early neonatal mortality rates, 1977

Number of units with paediatric staff per 100 low weight births	Number of units with 'in house' early neonatal mortality rates of:			
	<7	*7–9*	*>9*	*All mortality rates*
≤21	–	3	1	4
22–30	1	6	2	9
>30	4	–	–	4

Table 6 West Midlands maternity units grouped by paediatric staffing ratios and 'in-house' early neonatal mortality rates, 1982

Number of units with paediatric staff per 100 low weight births	Number of units with 'in house' early neonatal mortality rates of:			
	<7	*7–9*	*>9*	*All mortality rates*
≤21	1	1	3	5
22–30	5	1	–	6
>30	8	–	–	8

Summary
Very low birthweight and paediatric staffing are separately related to the risk of perinatal death. In most years studied there was a strong significant correlation between perinatal (PNMR) or first week deaths, and the proportion of very low weight births – this is to be expected and is seen in Tables 1 to 4. Mortality was also negatively correlated with measures of staffing and rates were lower when staffing ratios were higher. This correlation was usually very weak except for the correlation between mortality and number of paediatricians per low weight births and is seen for two years,1977 and 1982, in Tables 5 and 6. Units with high paediatric staffing ratios are unlikely to have high mortality and units with higher mortality do not have higher staffing ratios. Variation in perinatal mortality is primarily associated with proportion of births under 1500 g but higher paediatric staffing levels are associated with lower mortality rates.

Quality appraisal
1 Case mix adjustment
2 Other adjustment
3 Uniform data collection
4 Participant follow-up
5 Random sampling
6 Geographical dispersal

1 No case-mix adjustment reported. Clinical risk was estimated using the birthweight distribution and single/multiple birth.
2 Statistical analyses used regression to explain the health outcomes. Explanatory variables were staffing measures and birthweight distribution, dependent variables were measures of mortality. The analysis were sequentially adjusted for multiple births; congenital malformations; first week deaths; stillbirths; stillbirths and congenital malformations combined. Also studied was the effect of variation in ratio of qualified midwives alone (SCM) as opposed to midwives + nurses (MN), and the proportion of paediatricians to low weight births (PD/LBW) as paediatricians spend more time with high-risk births. Other allowances included the use of the annual number of births at each unit as a weighting factor in all the analysis.
3 Some gaps in data collection acknowledged and missing units were excluded from the analysis. HAA data has problems with coding errors and numbers of births for the region were liable to be incorrectly estimated. Staffing estimates also differed slightly from the regional manpower statistics.
4 Hospitals were the unit of analysis and data for each maternity unit was examined over six years.
5 Non-random sampling of data
6 20 units in West Midlands, UK

Commentary

The authors acknowledge there are shortcomings in the data, in their availability and completeness. Also the measure of mortality as an indicator may be open to question as the incidence of mortality now, and at the time of study, is very low. Wide variations in the ratios of professional staff and babies at each unit can exist. Selective transferrals and transfers between hospitals may also affect the interpretation of the findings. The study does produce useful insights into the effects of specialist staffing levels on health outcomes.

Research implications

The study has implications for staffing policy, but future research could focus on the use of suitable indicators to use across different units for determining advisable staffing levels. Routine maternity services monitoring should include audits using these agreed morbidity indicators and include client satisfaction to help decide the best use of staffing resources.

Table A2.6 Availability

ID, origin, authors (year)	12, USA, Arbabi, A., Jurkovich, G.J., Rivara, F.P. *et al.* (2003)			
Aims	To explore the effects of an in-house attending surgeon on-call policy and the presence of trauma and critical care fellowship programmes on critically injured patients *Workforce:* Surgeons, Level I trauma hospitals *Feature:* HR; in-house attending surgeon on-call policy (IH Policy = attending general surgeon remains in hospital 24 hours a day; no IH policy = attending trauma surgeon to direct the care of all trauma patients and be present in the hospital within 15 minutes for critically injured patients) and postgraduate fellowship programs in trauma and surgical critical care policy *Outcome:* Mortality, hospital length of stay (LOS) and intensive care unit (ICU) LOS			
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, two cohort studies 2 Inclusion for blunt injury cohort: discharged from participating institutions, at least 18, head abbreviated injury score of 2+, and fracture of at least lower extremity long bone (tibia and/or femur). Exclusion criteria were pregnancy, burn injury, spinal cord injury with paralysis, and transfer from another institution more than 24 hours after injury. Inclusion criteria for penetrating injury cohort: discharged from participating institutions, age 12+ and penetrating abdominal injury, excluding patients with any other body injury with abbreviate injury score >2, and pregnant or burned patients. 3 31 blunt trauma hospitals with a total of 601 patients and 24 penetrating trauma with 503 patients 4 In-hospital 5 Data were collected by medical record abstraction at each hospital and then collated by UHC. Information regarding trauma centre policy mandating was obtained by direct contact with trauma co-ordinators at the institutions. Blunt cohort: 1 June 1 1998 to 31 December 1998 and penetrating cohort: 1 November 1997 to 31 July 31 1998.			
Results Quantitative results	The presence of an IH policy (vs. no IH policy) had no effect on risk of fatal outcome in either cohort. However, presence of a trauma and critical care fellowship significantly decreased the risk of fatal outcome in the blunt trauma cohort after adjustments. IH policy had no effect on LOS in either cohort. However, presence of a trauma and critical care fellowship demonstrated a decrease in hospital LOS and ICU LOS in the blunt trauma cohort, but not for the penetrating cohort.			
			IH policy vs. no IH policy	**Fellowship program vs. no programme**
	Blunt	Fatal outcome	OR = 1.2 (0.5 – 3.0)	OR = 0.4 (0.1 – 0.8)
		Hospital LOS	Difference = –1.0 (3.8 – 2.0)	Difference = –3.2 (–5.9 - -0.6)
		ICU LOS	Difference = 1.4 (–2.0 – 4.8)	Difference = –4.7 (–8.8 - -0.6)
	Penetrating	Fatal outcome	OR = 1.7 (0.6 – 4.5)	OR = 0.9 (0.3 – 2.3)
		Hospital LOS	Difference = 1.7 (–0.3 – 3.2)	Difference = -0.2 (–1.6 – 1.2)
		ICU LOS	Difference = 0.12 (–2.0 – 2.2)	Difference = 0.8 (–1.4 – 2.9)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Injury Severity Score (ISS, 0–9, 10–15, 16–25 and ≥26), systolic blood pressure in the emergency department (0–59, 60–89 and ≥90 mm Hg), Glasgow Coma Scale (GCS) Score (<6, 6–8, 9–12 and 13–15) and age (<36, 36–55 and ≥56) 2 Hospital volume (high = ≥470 patients) 3 Yes 4 Unsure 5 Based on trauma centres that participated in the University Health System Consortium (UHC) Trauma Benchmarking Study. 6 Located throughout the United States – no specific details given
Commentary	The authors propose that the presence of a trauma and surgical critical care fellowship programme may be a marker for a mature, dedicated trauma and critical care service. The amount of data collected is small for the number of centres studied. In each cohort in each trauma centre, there are on average <20 patients. The injury matching is balanced by ISS, but the Abbreviated Score used between centres is not standardised, which introduces variability. The authors assume that self-designation of in-house vs. out-of-house leads to an actual difference in surgeon response. This was not measured or controlled for. It is possible that attending physician arrival was rapid for all severely injured patients in both groups. The composition of resuscitation teams varies in hospitals. Involvement of an experienced attending emergency department physician in an out-of-house hospital might balance the effect on care of an in-house hospital. In high-volume hospitals with experienced residents, the involvement of the attending physician will be less important, but one cannot make that conclusion in a hospital where volume is modest and where attending physicians are not needed.
Research implications	Does direct involvement of an experienced trauma surgeon improve outcomes? Does an IH policy translate into faster and more frequent attending surgeon presence at the bedside? A well-designed appropriately controlled study is needed.

ID, origin, authors (year)	659, USA, Barone, J.E., Ryan, C., Cayten, G. *et al.* (1993)
Aims	To determine whether patients treated at an institution desiring Level II trauma centre designation in a geographic area with a low incidence of penetrating trauma suffered any adverse effects because of lack of a 24-hour in-house OR staff *Workforce:* Operating room staff, teaching hospital vs. three Level I trauma centres *Feature:* Availability; control: surgeons and OR personnel in house at all times; case: surgeons in-house 24 hours a day but OR personnel on call at night (days: nursing staff were present from 7 am to 11 pm weekdays and 7 am to 3 pm on Saturdays; nights: the remaining 80 hours per week) and on weekends. *Outcome:* In-hospital mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, case–control 2 *Cases:* All patients who underwent surgery in the OR within 12 hours of admission were included. Patients dead on arrival or requiring Emergency Department (ED) thoracotomy and non-surgical cases were excluded. *Controls:* Patients were included if they were 13 or over, and either died or remained hospitalised for at least 48 hours. 3 Cases: 305-bed hospital, 659 patients 4 In-hospital 5 Data concerning major trauma patients were collected from the trauma registry by a single nurse-researcher at the study hospital from 1 July 1987 to 31 October 1991. Data concerning major trauma admissions at the three control centres were collected from ED logs from 1 July 1987 to 30 June 1989. Trained nurse-abstractors gathered the information using all available pre-hospital and hospital records.
Results Quantitative results	Survival probabilities were calculated for each Stamford Hospital patient and 4 possibly preventable deaths occurred. The lack of 24-hour in-house OR staff appeared to have no impact on the outcomes of these 4 patients. The z score determines whether differences between the number of observed deaths within any given group of patients is significantly different from the number predicted. An M score of at least 0.88 indicates a reasonable match with respect to the mix of injury severity between the test data set and the Major Trauma Centre Outcome Study (MTOS).

Mechanism	**Time**	**Number of patients**	**Deaths**	**Probability of survival**	**Z Score**	**M Score**
The Stamford Hospital (not designated or verified as a trauma centre)						
Blunt	Day	8	3	0.86 ± 0.15	–1.33	0.65
	Night	6	1	0.76 ± 0.39		
Penetrating	Day	3	0	0.92 ± 0.12	–1.26	0.52
	Night	5	2	0.67 ± 0.37		
Control						
Blunt	Day	22	7	0.65 ± 0.40	–0.14	0.50
	Night	14	7	0.60 ± 0.44		
Penetrating	Day	24	5	0.72 ± 0.41	0.93	0.34
	Night	33	4	0.89 ± 0.33		

Quality appraisal	
1 Case mix adjustment	1 TRIS methodology – ISS, RTS and age of each patient combined into a probability of survival value
2 Other adjustment	2 No
3 Uniform data collection	3 Yes
4 Participant follow-up	4 Not stated
5 Random sampling	5 No
6 Geographical dispersal	6 One suburban community teaching hospital and three trauma centres
Commentary	The time frames for the comparison groups were different. The volume of the hospital is extremely low: approximately 1.65 patients every month requiring surgical intervention in the first 12 hours, either day or night.
Research implications	A well-designed appropriately controlled study is needed.

ID, origin, authors (year)	355, USA, Doolin, E.J., Browne, A.M. and DiScala, C. (1999)
Aims	To analyse the individual components of paediatric trauma centres for their effect on patient outcomes *Workforce:* Surgeons and paediatric emergency department physicians, trauma centre *Feature:* Availability (in-house 24-hour presence of a surgeon) and presence of a paediatric ED physician *Outcome:* Length of stay (LOS) in paediatric ICU (PICU), mortality rate and overall LOS
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, cohort 2 Patients in the NPTR phase II (ending 1996) were the study group. 3 59 paediatric trauma centres 4 In-hospital 5 Each centre was asked to fill in a questionnaire regarding 8 components of a trauma centre with a dichotomous answer. The database of the NPTR was used to measure outcomes.
Results Quantitative results	The presence/absence of a surgeon or physician did not have any effect on the LOS in the PICU or on overall LOS. The presence of an in-house attending surgeon only reduced the mortality rate of the severely injured (ISS >35) older (>7 years) patient from 56.8% to 46.7% (p <0.05).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Patients were stratified to allow comparison: age ≥ or < 7 years and severity of injury ISS 1–16, 17–35 or >35. 2 None 3 Yes 4 15 centres were not included in the analyses as: inability to complete information or submission of <25 patients through phase 2. 5 Trauma centres that belonged to the National Pediatric Trauma Register (NPTR) 6 Not stated
Commentary	Unsure of how the authors arrived at the age categorisation, as other cut-off points have been suggested in the past research. Most adult trauma centres admit patients 16 and 17 years old and include them in their data. Hence it might have been better to stratify into different age groups. The authors did not report the average number of patients per centre per year for the study centres or if they looked at penetrating versus blunt trauma separately. Consequently, it is possible that an urban centre would see more penetrating trauma and that might have a higher impact on mortality rate, based purely on mechanism. It is recorded that the LOS of some children over 7 was 703 days and this appears to be an extremely lengthy stay for one patient. *P*-values were only quoted for significant results.
Research implications	What combinations (numbers, skill, grade or experience) of staff are required for optimal outcomes for patients?

ID, origin, authors (year)	942, USA, Pronovost, P.J., Angus, D.C., Dorman, T. *et al.* (2002)
Aims	To evaluate the association between ICU physician staffing and patient outcomes *Workforce*: Physician, ICU *Feature*: staffing intensity: High intensity staffing: 1) Closed ICU (primary physician is intensivist (critical care physician)), 2) Mandatory (no primary intensivist, but consultation is mandatory) Low intensity staffing: 3) Elective (intensivist consulted at request of physician), 4) No intensivist (intensivists were unavailable) *Outcomes* - Hospital and ICU mortality and length of stay (LOS)
Methods 1 Design 2 In-/exclusion 3 Sample size	1 Systematic review: randomised clinical trial (0); cohort study, historical control (19), concurrent control (2), both (1); case–control (0); cross-sectional, concurrent control (5) 2 Inclusion: randomised or observational controlled trials of critically ill patients (adults and children); ICU physician staffing strategies. ICU and hospital mortality and LOS. 3 ICUs = 156; patients (high intensity) = 14,356; patients (low intensity) = 13,117
Sources searched	Medline (1965–2001), EMBASE, Health Star and HSRPROJ via internet Grateful Med and The Cochrane Library (1998, issue 3). PubMed: Related articles feature. Hand search: annual scientific assemblies of the Society of Critical Care Medicine, the American Thoracic Society (1994–2001).
Validity criteria for primary studies	Risk of bias caused by temporal trends in mortality rates: low <2 years (15); medium 2–4years (8), high >4years (1) Risk of bias from confounding: low = used validated physiologic method for risk adjustment (21); medium = used selected clinical data (6); high = no risk adjustment (0) Risk of bias from incomplete follow-up: low 90% to 100% (27); medium 80% to 89% (0); high <80% (0)
Method of combining primary studies Investigation of differences and bias	Mortality rates were pooled using a random effects model. Length of stay was not pooled, but results displayed using a L'Abbe plot. Performed qualitative and quantitative assessment of heterogeneity. Publication bias investigated using funnel plot.
Results	High-intensity ICU phyisician staffing led to significant reductions in ICU and hospital mortality and LOS. High- vs. low-intensity ICU physician staffing: hospital mortality, pooled unadjusted RR = 0.71 (95% CI, 0.62–0.82). ICU mortality: pooled unadjusted RR = 0.61 (95% CI, 0.50–0.75). 10 out of 13 studies (77%) reported a reduction in LOS with high-intensity physician staffing (range of relative reduction: 5% to 42%). Pooled results in presence of qualitative heterogeneity. No quantitative heterogeneity found. Reported no evidence of publication bias.
Commentary	Two independent reviewers, third to solve discrepancies. All reviewers are intensivists, possible bias. Excluded non-English language papers. No studies followed up after hospital discharge – possible bias as early discharge may appear to reduce mortality. MeSH terms failed to identify all relevant articles in the search of databases. Other ICU characteristics not reported, such as nurse–patient ratio – staff intensity may be a proxy for another variable. Detailed tables of individual studies included in the paper. Funnel plot appears asymmetrical; however, authors report there to be no evidence of publication bias. No test has been performed to investigate this further. Quality of data source not reported. Pooled odds ratios are unadjusted; this may under-/overestimate the effect size.
Research implications	What are the features of high-intensity ICU staffing that effect outcomes? For example, nurse:patient ratio.

Table A2.7 Addition of a pharmacist

ID, origin, authors (year)	658, USA, Bjornson, D.C., Hiner, W.O., Potyk, R.P. *et al.* (1993)
Aims	To study the effect of pharmacists on health care outcomes *Workforce:* Pharmacists, Army Medical centre *Feature:* Additional team member of staff physician, physician resident, two physician interns and one medical student. *Intervention:* Two medical teams with a pharmacist (MTP) and one surgical team with pharmacist (STP) *Control:* Three medical teams without a pharmacist (MT) and two surgical teams without a pharmacist (ST)) *HO Groups (haematology-oncology program):* One pharmacist for inpatients, one for outpatients, one for nutritional support and drug information and one for internal medicine. *Outcome:* Morbidity (measured by LOS) and mortality. Cost-effectiveness was also measured but the results are not reported here.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Quasi-experimental, controlled trial 2 All general medicine and surgery patients admitted to the hospital. Patients who were transferred to or from a service cared for by a study team were excluded. Pharmacists were randomly assigned to two of the five general medicine teams and one of the three general surgery teams. 3 3 pharmacists, 3638 patients 4 30 days post-discharge 5 Classified each intervention as add drug, delete drug, change drug, change dosage, change route, provide pharmacokinetics consultation, educate prescriber, add order for laboratory test, or delete order for test. Whether each suggestion was accepted or rejected. Interventions were documented on patient-specific cards. Number of laboratory and radiologic procedures per patient, patient-specific nursing acuity scores, whether patients received discharge counselling by a pharmacist, the number of medications per patient at admission and at discharge, whether vaccinations were given to patients, whether an adverse drug reaction (ADR) was documented on the ADR reporting form, whether transfer to an intensive care unit occurred, and whether hospital re-admission occurred within 30 days after discharge. Data collection: 1 October 1990 to 30 September 1991 for medical patients and 1 February 1991 to 31 January 1992 for surgical patients.
Results Quantitative results	ANOVA revealed a significant difference in log LOS between the intervention and control groups.

Source of variation	**Sum of squares**	***DF***	**Mean square**	***F***	***P***
Log length of stay					
Team (medical or surgical)	0.000	1	0.000	0.002	0.965
Pharmacist (presence/absence)	0.736	1	0.736	4.599	0.032
Team–pharmacist interaction	0.038	1	0.038	0.237	0.626
Residual	492.140	3077	0.160		
Total	492.915	3080	0.160		

The chi-squared test showed no significant difference in mortality (X^2 = 1.68, p = 0.2) between the intervention (21 deaths, 1.75% mortality) and control group (46 deaths, 2.45% mortality). Most deaths (93%) occurred among patients on the medical wards. The percentage of patients documented to have experienced an ADR was greater in the intervention group (1.7%) than in the control group (0.5%) however no significance test was reported. Subgroup analysis showed no difference in the log LOS except when patients with an LOS of <=30 days were the only ones counted (intervention vs. control p = 0.015 and MTP vs. MT p = 0.045).

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 *Patient acuity:* composite average of vital signs, intensity of nursing monitoring, the patients' activities of daily living, types of feeding, need for and type of IV therapy, treatments, procedures, medications, need for and type of respiratory therapy, teaching and emotional support. Patients received weighted scores for each indicator and these were categorised as a score of 1 to 6 (higher scores = more nursing care). *Type and severity of illness:* recorded diagnosis-related group and major diagnostic category. The severity was rated with TOTSCALE (overall measure of resource consumption and disease severity) using ICD-9-CM diagnosis and procedure codes, age, sex, discharge status, admission source and DRG. 2 None stated 3 Yes 4 1560 patients excluded because they were either transferred to a study ward (624) or from a study ward (936). 5 No 6 One Army Medical centre in Washington
Commentary	Both pharmacists on the medical teams had post-baccalaureate Pharm.D. degrees (one with 20 years' experience and the other 5) and the pharmacist on the surgical team had entry-level Pharm.D. degree (one year practical experience). Assignment of patients to the groups was not randomised. However, comparison of the groups' age, sex, process-oriented variables, major diagnostic categories and related groups were well matched. The authors report that re-admission rates were similar for the groups, and therefore they can be more certain that the shorter length of stay in the intervention group was not associated with premature discharge. A multi-centre RCT with a large number of pharmacists, in which the pharmacists, rather than the patients, would be the unit of analysis would provide more information. Physician members of the team rotated to other teams monthly, so the influence of retention of the pharmacists may have underestimated the differences in outcomes.
Research implications	Does the training or experience of the pharmacist influence the outcomes? Is it that the pharmacist is just increasing the number of staff on the ward and therefore improving outcomes? Is it the mix of staff in the team that influences outcomes? Is it the grade mix of the team that affects outcomes? What is the pharmacist's role, what do they do and how do they do it? Would the same results hold if another staff member were added to the team? Were the re-admissions drug or disease related? Do pharmacists prevent transfer to more costly and more acute care wards? Do the results hold outside the study centre?

ID, origin, authors (year)	256, USA, Boyko, W.L., Yurkowski, P.J., Ivey, M.F. *et al.* (1997)
Aims	To investigate the influence of pharmacist's participation on economic and morbidity outcomes *Workforce:* Pharmacist (clinically trained with a Pharm.D. degree and pharmacy practice residency experience), tertiary care teaching Hospital *Feature:* Additional team member (team = attending physician, senior and junior medical residents and medical students) *Outcome:* LOS
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Quasi-experimental, controlled trial 2 Patients were included in the analysis if they were directly admitted to either the study team or the control team and stayed in hospital <21 days. Patients were excluded from the analysis if they were transferred into or out of either the treatment or control team, died, left the hospital against medical advice, or had received care from either team before the beginning of the study. 3 700-bed centre with >100 full-time equivalent employees in pharmaceutical care; 867 patients included 4 In-hospital 5 A specific staff pharmacist used a team consensus form to collect data for the control team patients. Information on new admissions to the control team and on the status of previously admitted patients was obtained bi-weekly in interviews with the senior medical resident. Patients were enrolled from August 1994 through April 1995.
Results Quantitative results	The addition of a pharmacist on an internal medicine team resulted in a significant reduction in LOS. **Outcome variable** / **Treatment group** / **Control group** / **Difference** / ***P*** LOS per admission (days) / 4.2 / 5.5 / 1.3 / <0.0001
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 None stated; no significant differences between groups in age and major diagnostic category, but more men in the treatment group. 2 None stated 3 Yes 4 From the 1180 eligible patients, 867 patients met the criteria for inclusion (414 in the treatment and 453 in the control). 313 patients were excluded: left against medical advice (treatment = 3, control = 4), Died (9, 9); stayed >=21 days (4, 8); cared for by the medical team before the start of the study (12, 14); transferred from another service (81, 69); and transferred to another service (51, 49). 5 Study patients were assigned to the treatment or the control team according to which team was on call at the time of admission. 6 One unit in Cincinnati
Commentary	Patients were assigned to groups in an unorthodox manner. However, the authors report that the two groups were well matched demographically and diagnostically. A severity index was unavailable in the hospital at the time of study. There were occasions during the study on which physicians (non-attending) participated on different teams from those to which they were initially assigned and hence the findings could have underestimated the results. Since the same individual served as the treatment team pharmacist throughout the study, the results may not be generalisable to the profession as a whole.
Research implications	Does the experience of the pharmacist matter? Does the presence of a pharmacist just add to the number of staff present or is it a particular additional role that the pharmacist provides? Do the results differ at weekends as compared to weekdays?

ID, origin, authors (year)	707, USA, Carter, B.L., Barnette, D.J., Chrischilles, E. *et al.* (1997)
Aims	To evaluate the impact of pharmacists on blood pressure control, quality of life (QoL), patient satisfaction, quality of care and cost of care *Workforce:* Community pharmacist, group medical practice *Feature:* Additional team member *Outcome:* QoL and patient satisfaction. (Blood pressure control, quality of prescribing and economic outcomes were evaluated but will not be reported here)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Quasi-experimental, controlled trial 2 *Inclusion:* Men and women of any racial group, 18+, with essential hypertension, who signed a consent form were eligible. All subjects had to be receiving care from one of the physicians in the medical centre or the annex, and prescriptions from the clinic pharmacy. They also had to have hypertension defined by predefined criteria. *Exclusion:* Secondary causes of hypertension; were unwilling or unable to return to the clinic pharmacy for scheduled appointments; had a spouse or sibling enrolled in the study; had blood pressure exceeding either 210mm Hg systolic or 115mm Hg diastolic; or had a serious complicating disease that was so disabling that blood pressure control was a secondary minor consideration. 3 29 patients were enrolled in the study groups and 26 in the control group (a sample size calculation was performed and 50 patients were needed). 4 Six months 5 Both groups completed the SF-36 (a short form measure of generic health status) at baseline and at the end of the study. Patients also received a questionnaire at the end of the study to assess their overall satisfaction with the delivery of care and with pharmacy services.
Results Quantitative results	*Quality of life:* At baseline the study group had worse QoL measures than the control group, with poorer scores in each of the eight categories. The only significant difference was in the bodily pain domain (p <0.016). After six months, the QoL scores increased in the study group and thrree of these increases were statistically significant. The control group had no significant changes in scores, but several categories showed slight declines.

Domain	**Control group at baseline**	**Control group at 6 months**	**Study group at baseline**	**Study group at 6 months**
Health perception	61.2	64.0	58.2	58.7
Physical functioning	66.5	67.7	61.5	70.7*
Role limitations, physical	63.5	62.5	54.3	74.0*
Role limitations, emotional	69.4	65.3	50.0	63.9
Social functioning	79.3	84.1	73.4	76.6
Mental health	75.5	75.7	73.4	71.0
Bodily pain	76.7	74.7	58.4**	71.1†
Energy, fatigue	55.0	56.3	47.5	54.1

* *significantly higher than the study group at baseline based on* t-*test of gain scores,* p <*0.05*
** *significantly lower than the control group at baseline based on* t-*test,* p <*0.01*
† *significantly higher than baseline in the study group based on* t-*tests of gain scores*

Patient satisfaction: patients in both groups were satisfied with the pharmacy services; 5 of the 13 questions were statistically significant at the 0.05% level.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 None stated 2 None stated 3 Yes 4 Two patients dropped out after the first month and upon review two patients didn't meet the in-/exclusion criteria. 5 Physical location dictated whether a patient was in the study or control group, but then patients were randomly chosen from the two groups using a table of random numbers. 6 Taylorville, Illinois (population approximately 10,000)
Commentary	The pharmacists had practised in this setting for 13 and 18 years, just not as part of the team with the same responsibilities, but were familiar with the surroundings. Pharmacists were given training prior to the study. The satisfaction survey used was generated for this study and has not been previously validated. Comorbid conditions were uneven with the study patients being less healthy, but didn't adjust for this. This study was conducted in one setting; ability to extrapolate the results to other settings is limited. The patients were required to fill out the satisfaction questionnaires in the pharmacy and consequently this may have resulted in higher satisfaction scores.
Research implications	Does the experience of the pharmacist matter? Did the fact that the pharmacists were familiar with the surroundings and of members of the workforce influence the results? Does the experience/grade of the other team members influence the results? Could another member of staff perform this role with the same specific training in hypertension? Additional investigation with multidisciplinary teams in private and rural practices should be conducted to determine the impact on patients seen in primary care.

ID, origin, authors (year)	350, USA, Leape, L.L., Cullen, D.J., Clapp, M.D. *et al.* (1999)
Aims	To measure the effect of pharmacist participation on medical rounds in the ICU on the rate of preventable adverse drug events (ADEs) caused by ordering errors. *Workforce:* Pharmacist, intensive care unit (ICU) and coronary care unit (CCU) in a teaching hospital *Feature:* Additional team member; participation on physician rounds *Outcome:* Adverse drug events (ADE) *Intervention:* Assignment of an experienced senior pharmacist to make rounds with the residents, nurses and attending staff each morning; was present in the unit for consultation and assistance to the nursing staff during the rest of the morning and was available on call as necessary throughout the day. The total commitment was approximately half the pharmacist's time. In the control ICU, as is the usual practice, another pharmacist was available in the unit for part of the day but did not make rounds.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Quasi-experimental, before and after comparison between Phase 1 (pre-intervention) and Phase 2 (post-intervention) and Phase 2 comparison with CCU. 2 Admission to the study unit during the study period. 3 17-bed medical ICU (control) and 15-bed CCU 4 In-hospital 5 Trained and experienced investigators (one nurse and one pharmacist) identified incidents by review of medical records in which they examined all progress notes, orders, and laboratory results during the index admission. Incidents were evaluated independently by two physician reviewers who classified them according to whether or not an ADE or potential ADE was present. If consensus could not be reached a third reviewer evaluated the incident. All reviewers and investigators were blinded to patient group assignment. The pharmacist completed a report form for each intervention that could potentially lead to a change in orders, noting the date, drug, nature of order, the specific recommendation and whether or not the physician accepted it. Phase 1: 1 February 1993 to 31 July 1993 and Phase 2: 1 October 1994 to 7 July 1995. The intervention began in May 1994.
Results Quantitative results	In the before-and-after comparison, the rate of preventable ordering ADEs per 1000 patient-days decreased in the study unit by 66% from Phase 1 to 2. When the intervention unit was compared with the control unit during the same time period (Phase 2), the rate of preventable ordering ADEs in the study unit was 72% lower than in the control unit. The preventable ordering ADE rate in the control unit rose slightly from Phase 1 to Phase 2, but this change was not significant (p = 0.76). When results were calculated in terms of number of patients (admissions), the differences in rates were similar: study unit, 12% in Phase 1 to 4% in Phase 2, and for the control unit 10% to 11%. The rate of all ADEs also decreased substantially in the study unit from Phase 1 to Phase 2. However, the rate rose in the control unit by 34.3%. During Phase 2 a total of 398 pharmacist interventions were recorded. Of these 366 were related to ordering.

	Study unit		**Control unit**	
	Phase 1	***Phase 2***	***Phase 1***	***Phase 2***
Average daily census	13.9	12.4	12.9	11.9
Total patient-days (n)	1061	861	461	644
Patients (n)	75	75	50	75
All adverse drug events (n)	35	10	16	30
Rate per 1000 patient-days	33.0 (27–39)	11.6 (8–15)*	34.7 (26–43)	46.6 (38–55)
Preventable ordering ADE (n)	11	3	5	8
Rate per 1000 patient-days	10.4 (7–14)	3.5 (1–5)*	10.9 (6-16)	12.4 (8–17)

*p <0.001

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 None stated 2 None stated 3 Yes 4 Not stated 5 Random number generator to select 75 patients from each of the 3 groups and to select 50 patients from all those admitted to CCU during Phase 1 to detect whether unmeasured variables may have altered the rate of ADEs. 6 1 tertiary care hospital in Boston
Commentary	Nurse and physician staffing ratios were similar for both units. Patients in the medical ICU had a range of acute and chronic medical illness other than primary cardiac disease, while those in the CCU were primarily cardiac patients. The authors only looked at one ICU in one teaching hospital and ADE are more common in teaching hospitals than in community hospitals and occur more frequently in ICUs, so these findings are not generalisable to all types of units or all types of hospitals. The results do not represent the full extent of preventable ADEs, since record review does not capture all events, nor does it capture most potential ADEs, the 'near misses' because they are seldom recorded in patients' charts. Physicians and nurses in this ICU function as a team and make rounds together. Pharmacists' participation would be more difficult to arrange in units where multiple physicians make rounds at different times. The success of this participation depends on interpersonal relationships. Thus, the personality and co-operativeness of the pharmacist and the medical staff are critical factors in making this system work.
Research implications	Does the introduction of a pharmacist on rounds reduce the workload of nurses? Is the addition of a pharmacist just increasing the workforce:patient ratios and therefore reducing errors? Does the experience or training of the pharmacist on the rounds matter?

Table A2.8 Substitution

ID, origin, authors (year)	310, USA, Aiken, L.H. *et al.* (1993)
Aims	Whether the primary care provided by a physician (MD) or a nurse practitioner (NP) influences HIV-infected patients' health outcomes *Workforce:* Physician (MD) or nurse practitioner (NP); primary care *Feature:* Substitution *Outcome:* Functional status, symptom occurrence, self management, health service use and patients' assessment of their care.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Observational retrospective cohort study 2 HIV/AIDS patients seen in the clinic at least once in the year prior to the current index visit to a hospital in Philadelphia. 3 87 patients (103 patients approached by researchers, participation rate 84%). 4 N/A 5 Self-administered questionnaire:N/A
Results Quantitative results	There is no evidence that outcomes of care differ substantially by provider type. In general patients were not dissatisfied with their care; The logistic regression reveals that NP patients reported 45% fewer problems with their care than MD patients (β = –0.595, p = 0.003); controlling for sex, NP patients reported more symptoms than MD patient (β = –0.235, p = 0.004); MD and NP patient were equally likely to engage in self-care action once they experience any of the nine symptoms. Two-thirds of all symptoms, on average, were improved by self-care actions, regardless of provider type; no significant differences by provider type were found for a variety of functional status measures; there were no significant differences in patterns of hospital admission, emergency room visits, or use of specialised mental health services between MD and NP patients.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 The two patient groups did not vary in terms of race, education, income or type of health insurance. However, the study only adjusted age in symptom occurrence analysis. Other risk factors, e.g. age and initial health status, were not adjusted. 2 N/A 3 Uniform 4 Complete 5 Convenience sampling 6 Single setting, one outpatient clinic of a university hospital
Commentary	Did not control for sex when measuring patients' satisfaction with the care. Other risk factors, e.g. age and initial health status, were not adjusted. Did not report the follow-up time and the data collection time. The study was conducted in a single site with a limited number of primary providers and used a convenience sample.
Research implications	More extensive use of nurse practitioners could potentially enhance access to care for persons with HIV-related illness wherever they live but particularly in already medically underserved areas. More robust controlled or observational studies on this area needed.

ID, origin, authors (year): 527, USA, Bolovinac *et al.* (1999)

Aims

To identify patient outcomes, as well as patient, RN, and UAP satisfaction levels following a change in patient care delivery using UAP as nurse extender.
Workforce: Registered nurses (RN), licensed practice nurses (LPN), and unlicensed assistant personnel (UAP)
Feature: Substitution: UPA substitute LPN: Period A: RN, LPN and UAP; Period B: RN and UAP
Outcome: Falls and post-implementation satisfaction

Methods

1 Design
2 In-/exclusion
3 Sample size
4 Follow-up time
5 Data collection: source and period

1 Prospective cohort study
2 *Included:* adult medical, surgical, post-angiography, post-cardiac catheterisation, blood transfusion and plasmapheresis recipients, inpatient chemotherapy, and wound care patients, who were discharged home from the short-stay medical–surgical unit discharged from 1 June 1997 to 21 July 21 1997.
Excluded: patients who were transferred to another unit, who were transferred onto the studied unit, and patients who remained in the unit for long-term care (more than 7 days).
All RNs who have worked using the UAP as RN extender care delivery system and who presently work on this unit were also included; all UAPs on the studies short-stay medical-surgical unit who have worked with the UAP as RN extender care delivery system and have completed orientation to this unit were also included.
3 40 patients, 15 RNs, and 9 UAPs.
4 Eight months for the sampled patients for patient satisfaction survey, four years for two independent samples for fall rates comparison.
5 For patient satisfaction: the self-administrated questionnaire (SAQ) completed by patients prior to discharge, during the first quarter of 1997 (Period A: patients cared by RN <LPN, and UAP) and the third quarter of 1997 (Period B: patients cared by RN and UAP)
For patient fall rates: quarterly reported patient falls records eight quarters before implementation and eight quarters after the implementation of the UAP patient care delivery system. But no mention on the implementation time of the UAP patient care delivery system.

Results

Quantitative results

One-sample t-test comparison of patient satisfaction surveys 1st quarter and 3rd quarter 1997 (33 patients finished the questionnaire)

Satisfaction with	**Mean**	**SD**	**Test statistic***	***t*-value**
Friendliness and courtesy of staff	90.91	13.72	73.20	7.42**
Skill, experience, and competency of staff	84.45	19.70	72.00	3.75**
Overall quality of care and services	83.33	20.41	72.70	2.99**
Staff listens to concerns and opinions	82.03	24.78	70.60	2.61***
Timeliness of assistance for personal needs	75.00	27.24	64.80	2.15***
Cleanliness and appearance of room	75.24	30.93	66.50	1.44 (NS)
Clear complete explanations	71.97	32.33	64.20	1.38 (NS)

* means 1st quarter patient satisfaction scores on the same items
** $p \leq 0.01$
*** $p \leq 0.05$
NS not significant

The patient fall rates of pre-, and post-implementation were compared using a *t*-test for two independent samples using the time periods as the independent variables and the number of falls as the dependent variable.

t-test for 2 independent sample comparison of falls 8 quarters prior to implementation of UAP patient care delivery system and 8 quarters following implementation

Period	**Number**	**Mean**	**SD**	**DF**	***t*-value**
Prior to implementation of UAP program	8	12.00	4.84	4	0.00 (NS)
Following implementation of UAP program	8	12.00	4.84	14	0.00 (NS)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 None 2 None 3 Uniform 4 33 patients completed the satisfaction questionnaire. 5 Convenience sample 6 One unit in one hospital
Commentary	No mention on the two independent samples for the comparison of the fall rates. It is not clear if the care model change (from RN, LPN, and UAP to RN and UAP) happened at the 2nd quarter of 1997 for measuring patient satisfaction is the implementation of the UAP patient care delivery system mentioned in comparing the patient fall rates. It is a single setting study with a small sample size.
Research implications	Additional research is needed on a larger scale, with a homogeneous patient population, to facilitate data-driven decisions on the efficiency of the UAP programmes. Additional data on a variety of clinical outcomes should be collected to determine if quality of care changed with the use of UAPs. The role of the RN in delegating responsibility to UAPs should also be examined to determine which method of task delegation is most efficient.

ID, origin, authors (year)	179, UK, Aubrey, W.R. and Yoxall, C.W. (2001)
Aims	To evaluate the effectiveness of advanced neonatal nurse practitioners (ANNPs) in resuscitation of pre-term babies at birth against the standard set by junior medical staff *Workforce:* Advanced Neonatal Nurse Practitioners (ANNPs); tertiary care *Feature:* Substitution of junior medical staff by ANNPs *Outcome:* Resuscitation details and other clinical outcomes
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Retrospective analysis 2 Babies born in Liverpool Hospital at <33 weeks gestation and needed resuscitation, between January 1998 and April 1999 3 256 babies met the inclusion criteria, 245 had a full data set available. 4 None 5 Data were from Liverpool Hospital records kept between January 1998 and April 1999; resuscitation details, temperature on admission to neonatal unit, basic data and clinical outcomes were recorded.
Results Quantitative results	ANNPs are effective in resuscitation of pre-term babies. ANNP-led teams were no more likely to intubate, but they intubated earlier and administered surfactant sooner. Babies in ANNP-led teams were less likely to be hypothermic on admission to the neonatal unit.

	ANNP-led teams	**Medically led teams**	***p*-value**
Number of infants	76	169	
Birth weight (g)	1242 (530–2200)	1242 (440–2440)	0.88
Gestation (weeks)	30 (24–32)	29 (23–32)	0.17
Cord pH	7.32 (6.8–7.46)	7.32 (6.7–7.47)	0.76
Apgar (1 min)	6 (0–9)	6 (1–10)	0.32
Apgar (5 min)	9 (0–10)	9 (0–10)	0.67
Caesarean section	53/76	84/169	0.005
Time to intubation	2 min (20 sec–10 min)	3 min (1–18min)	0.0001
Time to surfactant administration (min)	8 (3–20)	10 (2–150)	0.0005
Intubation attempts	1 (1–3)	1 (1–4)	0.91
Admission temperature <35 degrees C	2/61 (3%)	25/145 (17%)	0.013
Admission documentation completed	63/76 (82%)	113/169 (67%)	0.015

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 and 2 No 3 Yes 4 No 5 No 6 Liverpool Hospital
Commentary	At the time of this study ANNPs were only working weekday shifts; therefore their Caesareans were usually scheduled, whereas night-time and weekend Caesareans are usually emergency and performed with medical staff. This could have affected some of the data and biased the medical staff. Time to intubation results could be due to issues with timekeeping, or it could reflect that ANNPs have a higher awareness of recognising when infants need intubation.
Research implications	Study needs to be re-conducted on a larger scale and with an objective timekeeper. Need to look at patient satisfaction rates of parents. Study if the minor time differences in intubation and surfactant administration cause differences in any long-term health outcomes.

ID, origin, authors (year)	425, UK, Boulton, B.D., Bashir, Y., Ormerod, O.J.M. *et al.* (1997)
Aims	To establish the feasibility and safety of an appropriately trained clinical nurse specialist performing diagnostic cardiac catheterisation *Workforce:* Appropriately trained clinical nurse specialists (ATCNS), primary care *Feature:* Substitution of cardiology registrars for ATCNS in performing low-risk cardiac catheterisation *Outcome:* Procedural complications
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort 2 Patients undergoing non-emergency cardiac catheterisation for investigation of ischaemic heart disease. Patients undergoing left ventricular and native coronary angiography via the transfemoral approach for standard clinical indications. Patients requiring a transbrachial approach and those with valvar heart disease, congenital heart disease, or a history of coronary artery bypass surgery were excluded. High-risk cases such as those with ongoing myocardial ischaemia, hypotension/shock, pulmonary oedema, and uncontrolled arrhythmias were also excluded. Patients admitted with unstable angina that had settled with medical therapy before investigation were eligible for the study. 3 200 patients 4 None 5 Not specified how data was collected or time period.
Results Quantitative results	There was no significant difference between the ATCNSs and cardiology registrars. ATCNSs performed the procedure more quickly.

	Registrars (200 patients)	**Nurse** (first 100 patients)	**Nurse** (second 100 patients)	**Nurse** (overall)
Patient's age (mean (SD) years)	61.4 (10.3)	63.4 (9.2)	60.3 (15.3)	61.9 (12.7)
Male:female	133:67	75:25	74:26	149:51
Unstable angina (%)	44.0	51.0	47.0	48.5
Procedure duration (mean (SD) min)	33.1 (12.4)	30.1 (12.7)	30.3 (12.5)	30.2 (10.3)
Fluoroscopy time (mean (SD) min)	6.0 (3.8) $p = 0.03$	5.0 (3.4) $p = 0.9$	3.8 (2.3)	4.4 (2.9)
Complications	4/200 $p = 0.05$	1/100 $p = 0.004$	1/100	2/200

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Yes, by age, sex, and unstable angina 3 No mention of how data was collected 4 No 5 No 6 One cardiac center in UK
Commentary	Needs to be tested on a larger scale to protect against nurses' volunteer bias. Allowing specialist nurses to take over this duty allows cardiologists and other medical professionals to deal with more high-risk patients and focus on patient care rather than testing.
Research implications	Needs to be done on a larger scale. Need to look at patient satisfaction: do nurses bring a better set of intrapersonal skills?

ID, origin, authors (year)	21, UK, Caine, N., Sharples, L.D., Hollingworth, W. *et al.* (2002)
Aims	To assess the feasibility and safety of nurse practioner-led outpatient clinics and their acceptability to patients and their doctors. In addition to compare the cost-effectiveness of nurse practitioner-led care with a doctor-led system of care. *Workforce:* Nurse practitioners (NP), general practitioners (GP); primary care *Feature:* Substitution of GPs with NPs *Outcome:* Health status, quality of life, lung function (as measured by forced expiratory volume in 1 second [FEV1])
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Random crossover controlled trial 2 Patients had to be over 18 years old with moderate to severe bronchiectasis confirmed by high-resolution CT scan. Excluded if life expectancy is less than 2 years, had an expected need for transplantation listing within 2 years, and FEV1 value less than 30% of that predicted, any other significant pathology that would modify management of bronchiectasis. 3 80 patients 4 2 years of study, no follow-up after duration of study 5 2 years of collecting FEV1 measurements and consultation data, SF-36, the Chronic Respiratory Index Questionnaire (CRIQ), St George's Hospital Respiratory questionnaire
Results Quantitative results	*Clinical results* **Nurse-led care** / **Doctor-led care** / **Mean difference (nurse–doctor (95% CI))** FEV1 litres: 1.87 (0.78) / 1.86 (0.81) / 0.01 (–0.04 to 0.06) FEV1 (%): 69.7 (20.8) / 69.5 (21.7) / 0.2 (–1.6 to 2.0) FVC (%): 87.6 (19.3) / 87.6 (19.4) / –0.02 (–1.5 to 1.4) 12-minute walk distance (meters): 765 (188) / 746 (197) / 18 (–13 to 48) The only demonstrable difference in clinical outcomes was the number of hospital admissions. There were more patient admissions under nurse practitioner-led care, although the readmission rates for bronchiectasis-related problems were not significantly different.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Patients mixed for respiratory function 3 Yes, FEV1 measurements and consultation data, SF-36, CRIQ, St. George's Hospital Respiratory questionnaire 4 Not beyond 2 years of study 5 Yes, by patients' respiratory function, organised by the hospital's research and development unit 6 Papworth Hospital, UK
Commentary	Only done in one specialised unit in a hospital; not generalisable to less-specialised units, clinics, and other diseases. NPs had a better rate of antibiotic compliance than GPs.
Research implications	Does this substitution method work with other disease processes?

ID, origin, authors (year)	351, Australia, Chang, E., Daly, J., Hawkins, A. *et al.* (1999)
Aims	To investigate whether nurse practitioners are able to provide a level of primary health service applicable to remote/isolated settings in wound management and treatment of blunt limb trauma *Workforce:* Nurse practitioners (NP); primary care *Feature:* Substitution of NP to provide a level of primary health service applicable to rural/remote/isolated settings in the emergency department for wound management and treatment of blunt limb trauma. *Outcome:* Patient satisfaction; evaluation of the clinical outcome of the wounds. Methodologies for the evaluation of care provision and development of NP training programmes were studied but not in relation to patient outcomes.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Included all clients identified as potential study participants by the triage nurse, who presented to the emergency department with blunt limb trauma or open wounds to the scalp, lower leg or forearm, between 07:00 and 24:00 when there was a NP on duty. Categories included a range of problems and interventions including: blunt trauma; insect/animal bites; crush injuries; contaminated wounds; burns; simple fractures; lacerations and simple suturing; wound management; administration of local anaesthetic; administration of tetanus toxoid; prescription of limited antibiotics and pain relief greater than paracetamol. Excluded: children under 10; clients with significant presenting and continuing vital signs alterations; multiple trauma; high-risk mechanisms of injury; concurrent health problems in need of urgent treatment; resuscitation. 3 232 participants with open and closed wounds and/or blunt limb trauma. 63 were supervised cases in the pilot trial. 91 randomised to medical practitioners and 78 to NPs. 4 4-month supervised competency trial and a 3.5-month unsupervised comparative study 5 Telephone interviews; client records; invitation to a review by the consultant orthopaedic surgeon to evaluate individual wounds for cosmesis and function
Results Quantitative results	Multivariate analysis was carried out on the five-interval scales of measurement, and differences in these scores between the clients of the two groups tested. Overall there were no significant differences between the two groups in all areas of care. No significant difference in waiting time between the two groups. (Values not reported.) 16 follow-up wound assessments by an orthopaedic surgeon were conducted. The majority of outcomes were rated between 7 and 10 on both dimensions – values not reported. The study found strong support for the role of NP in rural emergency setting with service choice enhanced by their availability. NPs were accepted by medical staff and participants in the study. There was no significant difference in client satisfaction between the groups. NPs found job satisfying, rewarding and overall a worthwhile and positive experience. The findings suggest that provision of this service may have potential benefits for isolated areas.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 N/A 2 N/A 3 Yes 4 132 (78%) of those randomised were followed up for client satisfaction. Loss of follow-up: 7% changed address, 6% had no telephone, 8% uncontactable, 2% incorrect telephone numbers. 5 Yes 6 1 of 11 pilot sites located in differing clinical settings throughout metropolitan and rural New South Wales
Commentary	The small sample size in the study places limitations on the degree to which one can generalise from the results. In evaluating the study NPs had ready access to medical practitioners at all times during the project. This means their ability to perform at the level achieved in the study was not tested in remote/isolated areas. No-one left the study – some participants were reassigned in the study – this was justified on the basis of and to comply with national triage scale times.
Research implications	Further research is required to measure the efficacy of NPs utilising the selected competencies in remote/isolated settings.

ID, origin, authors (year)	895, Australia, Charles, A., Le Vasseur, S.A. and Castle, C. (1999)			
Aims	To investigate a programme enabling clinical nurse specialists to suture minor lacerations in the emergency department *Workforce:* Clinical nurse specialists (CNS) and nurse educators; tertiary care *Feature:* Substitution if CNS to suture minor lacerations in the emergency department (ED) *Outcome:* Patients' waiting time including total time until seen by CNS/doctor, total time until suturing, total time spent in ED, patients' perception of waiting time, patient satisfaction, wound complications and healing outcome			
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised control trial. 2 Inclusion criteria included: presenting to ED with simple lacerations, aged 16 years and older, no bony involvement or neurovascular damage to the area to be sutured, no facial or perineal areas, no pre-existing medical problems and informed written consent obtained. 3 80 patients were recruited. 11 CNS and 2 nurse educators underwent the education process and participated in the trial. 4 Not stated 5 Patient questionnaire; follow-up assessment of the wound			
Results Quantitative results		**Mean rank**	**Cases**	**Group**
	Total time until seen by	44.67	40	Medical
	CNS/doctor	36.33	40	CNS
				2-tailed P
				0.137
	Total time until suturing	43.88	40	Medical
	commenced	37.13	40	CNS
				2-tailed P
				0.194
	Total time spent in ED	44.67	40	Medical
		36.33	40	CNS
				2-tailed P
				0.108
	Patients' perception of waiting	40.05	40	Medical
	time in ED	39.95	39	CNS
				2-tailed P
				0.984
	Patients' perception of care	33.35	40	Medical
	received	47.65	40	CNS
				2-tailed P
				0.0016

	Adequate approximation of the wound — **Patient groups** (*CNS* / *Medical group*) — **Total** Yes: 35 / 37 — 72 No: 5 / 3 — 8 Total: 40 / 40 — 80 **Complications with wound** — **Patient groups** (*CNS* / *Medical group*) — **Total** Yes: 4 / 5 — 9 No: 36 / 35 — 71 Total: 40 / 40 — 80 Waiting times were not significantly different between patients sutured by doctors and by nurses. However, patients cared for by the CNS group appeared to be more satisfied with their care and the overall services received. Wound healing outcomes were found to be similar between two groups. This study demonstrates the new CNS role is capable of providing high-quality care to individuals with minor lacerations who typically wait to receive medical attention. The role also allows doctors to use clinical time more effectively in the medical management of seriously ill patients.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 N/A 2 N/A 3 Yes 4 Complete 5 Yes, by triage nurse 6 One medical centre in the south-eastern suburbs of Melbourne
Commentary	The study was open to the possibility of subversion since patients were randomised using unmarked envelopes opened by the triage nurse. The fact that waiting times were not significantly different between patients sutured by doctors and by CNS may reflect the inexperience of the nurses who had not sutured prior to the introduction of this programme or the need for a doctor to review patients seen by the triage nurse or sutured by the CNS.
Research implications	Could this programme be implemented in other emergency departments?

ID, origin, authors (year)	898, USA, Farr, G., River, R., and Amatya, R. (1998)
Aims	Two objectives: 1 To test if non-physicians, such as nurses, nurse practitioners, physician's assistants and midwives, properly trained and experienced in IUDs, could perform IUD insertions as safely as physicians by comparing rates of insertion failures and complications between these two groups 2 To assess the use–effectiveness of IUDs when inserted by physicians and non-physicians by comparing rates of continuation and termination due to expulsion to removals for pregnancy, medical or personal reasons *Workforce:* Physicians – ob/gyn physicians and general practitioners; non-physicians: medical students, and nurse and midwife *Feature:* Substitution – physician group and non-physician group *Outcome:* Rates of insertion failures and complications; rates of continuation and termination.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Randomised controlled trial 2 Healthy women who had no contraindications for IUD use, were sexually active, between 18 and 40 years of age, had a prior pregnancy, and had given informed consent to participate in the study. IUD insertions were performed during the interval period (last pregnancy to have ended at least 42 days prior to IUD insertion). Subjects were asked to return to the clinic at 1, 3, 6, and 12 months after IUD insertion and at any time complications occurred. Physical and pelvic examinations were performed during each clinical contact with the participant. Subjects were discontinued from the study if they became pregnant or if their IUD was expelled, displaced, or removed for any reason. 3 367 women at three sites (147 in Nigeria, 71 in Turkey, and 149 in Mexico); 193 in physician group, and 174 in non-physician group 4 1 year 5 Socio-demographic characteristics, reproductive and contraceptive histories, and pre-existing medical conditions: collected at study admission Events related to IUD insertion: admission record forms The occurrence of subsequent pertinent events: case report forms at each clinic visit
Results Quantitative results	The trained non-physician health care worker can provide IUD services as safely and effectively as physicians, although additional training in evaluating medical contraindications remains necessary.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patients' age, months since last pregnancy outcome, years of education, parity, percentage of different previous contraceptive methods in last month, the willingness to have additional children, marital status were not significantly different in each group in three study sites. 2 No 3 Uniform 4 One year 5 No 6 Three countries

Commentary	Did not report the exact time of the trial. No mention on the sampling method. The study did not record the information on the health care providers' previous level of experience in inserting IUDs, thus it was not able to assess the experience level as a possible co-factor for risk of expulsion.
Research implications	It is not clear why non-physician performed more pain-free insertions than physicians. One possible explanation is the time spent with patients by the providers. Non-physicians may have taken more time to explain the insertion procedure with each patient and may also have been more gentle in performing the IUD insertion. Observational studies on variations in provider services are needed to accurately assess this possibility. It is not clear why non-physician insertions resulted in an overall higher expulsion rate by month 12. One possible explanation may be that physicians in these sites may have been more experienced in IUD insertions than were the non-physician staff. This supports the need for appropriate training of non-physicians in IUD insertion techniques. Such training must be designed utilising a competency-based approach. It is not clear if the same physician/non-physician who inserted the IUD made the determination for recommending removal. More observational studies on provider services are needed. It is not clear why there is no difference in the incidence of urogenital infection between the two groups in the Turkey or Nigeria centres, but Mexico centre had over twice as many women with a non-physician insertion subsequently experiencing at least one urogenital infection than did women having a physician insertion. Further study may be warranted. Special attention at ensuring that the IUD is inserted in the uterine fundus may be crucial in helping to reduce the likelihood of expulsion among insertions performed by non-physicians. Expanding training in IUD service provision to non-physicians could result in a higher utilisation of the contraceptive method.

ID, origin, authors (year)	197, USA, Fong, N.I., Holtzman, S.R., Bettmann, M.A. *et al.* (2001)
Aims	To determine the natural history of and outcome involved with peripherally inserted central catheters (PICCs) placed at a single institution and examine potential differences in the natural history of PICCs placed by interventional radiologists versus registered nurses *Workforce:* Interventional radiologists (IR), Registered Nurses (RN), tertiary care teaching hospital *Feature:* Substitution *Outcome:* Complications associated with peripherally inserted central catheters (PICCs)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Observational 2 All patients receiving PICCs at the study centre over 6.5 months 3 256 patients 4 In-hospital 5 Patient information was obtained by chart review, telephone interviews with patients and contact with medical staff responsible for the patient's care between 14 June, 1999 and 31 January, 2000.
Results Quantitative results	Of the 130 PICCs placed by IRs, therapy was successfully completed with 68.3% (n=83) and premature removal was required for 30.8% (n=40). Of the 192 PICCs placed by RNs therapy was successfully completed with 72.9% (n=140) and premature removal was required for 23.4% (n=45). The only significant difference between RN and IR placement was in the rate of premature removal as a result of occlusion. Four patients developed other complications (RNs = 3, IR = 1). Infection rate for IRs was 18.27 per 10,000 PICC days and the rate for RNs was 8.3 per 10,000 PICC days. These differences were not statistically significant (p = 0.77). No patients died directly as a complication of their PICC.

Complication	**Premature removal rates (%)**			
	Overall (n)	***IR (n)***	***RN (n)***	**p-*value***
Occlusion	5.9 (19)	9.2 (12)	3.6 (7)	0.02
Suspected Infection	5.6 (18)	6.2 (8)	5.2 (10)	0.58
Phlebitis	4.7 (15)	3.1 (4)	5.7 (11)	0.34
Mechanical failure	3.7 (12)	4.6 (6)	3.1 (6)	0.50
Inadvertent removal	5.3 (17)	6.9 (9)	4.2 (8)	0.24

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 None stated. Investigated age, sex, immune status and significantly more females in the control group. 2 None stated 3 Yes 4 Complete follow-up was obtained for 308 PICCs; 14 patients were lost to follow-up (7 in each group), used ITT analysis 5 No 6 One teaching hospital
Commentary	The two groups were not equivalent. The IR group represented the subset of difficult or problematic PICC placements from the RN group. Evidence for selection bias was the significantly higher percentage of female patients in the IR group. IRs were essential for complicated placements, exchanges, advancements and repositioning of PICCs. No adjustments were made for patient or hospital characteristics. Patients were not randomised to the two groups. These results are based on patients in one teaching hospital with RNs who were trained to perform this role. Small sample size.
Research implications	Further investigation of this type of substitution with adjustments for case mix in multi-centres. If the nurses are performing new roles who will fill their old roles? Does the prior experience of the RN make a difference? Which type of patients can RNs treat?

ID, origin, authors (year)	224, USA, Freedman, B.M. and Earley, R. (2000)
Aims	To determine whether there were differences in outcome between patients treated by a trained physician and patients treated by a trained, supervised nurse *Workforce:* Trained supervised nurses; secondary care *Feature:* Substitution of physician by trained supervised nurse to remove unwanted hair using an alexandrite long-pulsed infrared system *Outcome:* Reduction in hair growth, patient satisfaction, transient skin changes
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Longitudinal study 2 Patients were examined by a physician to determine whether their excess body hair was caused by an underlying metabolic disorder 3 100 patients with unwanted body hair on face, torso and extremities 4 12 months 5 Photographs, patient survey, examination
Results Quantitative results	**Outcome** / **Group P N = 50 Physician-treated** / **Group N N = 50 Nurse-treated** Reduction in hair growth (physician-assessed): 74 (8%) Not significant 70 (6%) Reduction in hair growth (patient-estimated): 75 (7%) Not significant 75 (5%) Patient satisfaction *: 1.6 (0.3) Not significant 1.4 (0.3) * Using an assessment scale of 1 (excellent) to 5 (poor) **Side effect** / **Group P N = 50 Physician-treated** / **Group N N = 50 Nurse-treated** Hyperpigmentation: 2 / 2 Hypopigmentation: 3 / 3 Blistering: 2 / 2 Scabbing: 1 / 0 Total: 8 (16%) / 7 (14%) This study concluded that properly trained nurses can safely and effectively perform laser hair removal while assuring a level of care that satisfies both patient and medico-legal concerns.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 N/A 2 N/A 3 Yes 4 Complete 5 No 6 Not stated
Commentary	The study is limited by its design. A randomised control trial would have been a more appropriate method of determining the differences in outcomes between patients treated by physicians and patients treated by a trained supervised nurse.
Research implications	

ID, origin, authors (year)	1147, UK, Gallagher, M., Huddart, T. and Henderson, B. (1998)
Aims	To determine the impact of telephone triage, conducted by a practice nurse, on the management of same-day consultations in a general practice *Workforce:* Practice nurse; primary care *Feature:* Substitution and work load *Outcome:* Patient satisfaction, repeat consultations for same problem
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Before and after 2 Patients calling and requesting to see a GP on the same day at one practice 3 1263 consultations, 192/271 responded to patient satisfaction questionnaire 4 None, unless in need of a repeat consultation 5 Repeat consults data taken from both written and computer records of patients from August 1995 to October 1995. Patient satisfaction questionnaire was sent through the post in June 1996.
Results Quantitative results	Doctor workload fell by 54%, from 1522 to 664 consultations per 3-month period. 154 (88%) patients were very or fairly satisfied with nurse telephone advice. Only 10 (6%) were fairly or very dissatisfied with telephone advice from the nurse. *Repeat consultations for the same problem at one and four weeks*

	Nurse telephone advice only	**Nurse surgery**	**Doctor surgery**	**Nurse and doctor surgery appt.**
	Frequency (Group%) n=325	Frequency (Group%) n=273	Frequency (Group%) n=565	Frequency (Group%) n=99
Repeat consultations within 1 week	78 (24.0)	41 (15.0)	67 (11.9)	19 (19.2)
Repeat consultations after 1 week and up to 4 weeks	54 (16.6)	38 (13.9)	116 (20.5)	22 (22.2)
Total number of repeat consultations within 4 weeks	132 (40.6)	79 (28.9)	183 (32.4)	41 (41.4)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 and 2 No 3 Yes 4 Only if repeat consultation; some telephone-only patients contacted with questionnaire 5 No 6 One general practice in South Tyneside, UK
Commentary	Questionnaire needed to be sent closer to date of nurse advice being given. The 3-month time span was too short and differences in the drop in numbers of appointments could be attributed to a 'holiday period'.
Research implications	Needs to be studied over a longer period of time. Can more patients be triaged and avoid coming into the practice all together?

ID, origin, authors (year)	100, UK, Horrocks, S., Anderson, E. and Salisbury, C. (2002)
Aims	To determine whether nurse practitioners can provide care at first point of contact equivalent to doctors in primary care *Workforce:* Nurse practitioners; primary care (general practice, out-of-hours centres, walk-in centres and emergency departments) *Feature:* Substitution with doctors at first point of contact *Outcome:* Patient satisfaction, health status, rate of prescription, referrals
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 Inclusion: RCTs and observational studies with prospective experimental design; nurse practitioners and doctors working in a similar way as concurrent controls; nurse practitioner defined as providing care at first point of contact, making initial assessment and managing patients autonomously. Studies from developed countries (Europe, North America, Australasia, Israel, South Africa and Japan). Outcomes must include at least one of following; patient satisfaction, health status, health service costs, or process of care measures (consultation length, number of prescriptions, investigations, referrals, admissions, return consultations, patient adherence, or measures of quality of care). 3 – 4 RCTs (11): general practice setting (9), emergency department (2). Observational (23): general practice setting (17), emergency department (6). 5 Medline (1966–2001), Embase (1980–2001), CINAHL (1982–2001), science citation index, database of abstracts of reviews of effectiveness, national research register, Cochrane controlled trials register and the specialist register of trials maintained by the Cochrane Effective Practice and Organisation of Care Group. Educational centres offering training for nurse practitioners in the UK and nurse practitioner organisations in the USA, South Africa and Australia were contacted for unpublished studies. 6 Validity criteria for primary studies: Quality appraisal was based on the criteria of the review group of the Cochrane Effective Practice and Organisation of Care Group. Methodological quality of RCTs used the following criteria: Presence of allocation concealment, follow-up for 80% of participants in doctor and nurse practitioner arms, blind assessment of outcomes or objective measures, outcomes assessed at baseline, reliable outcome measures, allocation by practice or site to protect against contamination. 7 Meta-analysis of RCTs where at least two studies had data on a particular outcome. Findings from observational studies were compared qualitatively. Heterogeneity between studies was measured for: individual consultations and those investigating care over time; nurse practitioners with different levels of training; different settings. Sensitivity analysis used to explore impact of including/excluding studies based on factors listed above
Results Quantitative results	Availability of nurse practitioners led to higher patient satisfaction and high quality of care. Patient satisfaction: 5 RCTs using continuous data found greater satisfaction with nurse practitioners (SMD*: 0.27 (0.07 to 0.47); 3 RCTs using dichotomous data found no significant difference between the groups; overall effect $z = 0.85$, $p = 0.4$. Analyses were performed in the presence of heterogeneity. Health status: 7 RCTs showed differences between measures and episode of care length; hence not analysed with meta-analysis, but showed no significant difference in health outcomes between groups. Process measures: not reported in this abstraction Quality of care: not reported in this abstraction (*standardised mean difference)

Commentary	Two independent reviewers, third to solve discrepancies. Great variety in outcome measures between the studies, difficulty in measuring health outcome after single consultation. Studies were not powered to detect rare but serious adverse outcomes. Few RCTs identified, and not all of them recent. Observational studies were generally of poor quality. Considerable heterogeneity between studies for all factors investigated. Detailed tables of individual studies and quality assessments available from bmj.com. Majority of studies focus on same-day appointments for minor illnesses, small part of the doctors' role.
Research implications	Need for large study with adequate length of follow-up in order to achieve power to detect ability of GP/nurse practitioner in detecting rare outcomes. What are the factors that lead to satisfaction in care? Trials are needed that investigate nurse practitioners and doctors working under similar circumstances, e.g. same rates of booked consultations, similar pressures. Research needs to look at consultations with patients with chronic diseases and complex psychosocial problems. Different models of organisation need be investigated, such as several nurse practitioners providing care at first point of contact supported by smaller number of GPs. Definition of nurse practitioner is unclear in UK, need to study training and skills of these nurses that lead to benefits found.

ID, origin, authors (year)	281, Ireland, Murphy, A., Bury, G., Plunkett, P., *et al.* (1996)
Aims	To see whether care provided by general practitioners to non-emergency patients in an accident and emergency department differs significantly from care by usual accident and emergency in terms of process, outcome and comparative costs *Workforce:* General practitioners; primary care *Feature:* Substitution of Accident and Emergency staff for GPs *Outcomes:* Unplanned re-admissions within 30 days of first visit, patient satisfaction, health status after 30 days
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Patients in one A&E department of 'semiurgent' or 'delay acceptable' patients without referrals from their GPs 3 4684 patients 4 Follow-up completed by 258 patients 30 days from first visit. 5 Health status was measured by a 4-question survey administered by phone or letter 30 days after attending A&E. Hospital re-admission was measured by the hospital's mainframe computer. Patient satisfaction was measured directly after consultation by a blinded interviewer with a standard questionnaire. August 1993 to October 1994.
Results Quantitative results	*Re-admission:* 4601 patients total **Reattending within 30 days** — **Mean number of visits by patients reattending** GP: 393 (17%) — 1.6 A&E: 418 (18%) — 1.8 *Patient Satisfaction*: 435 patients completed the consultation satisfaction questionnaire (GP patients=276, AE patients=159) **Mean** / **Median** / **SD** General satisfaction: GP: 67.8 / 71.0 / 19.5 A&E: 67.0 / 67.0 / 20.8 Depth of relationship: GP: 48.0 / 50.0 / 17.6 A&E: 47.0 / 50.0 / 17.9 Perceived time: GP: 55.8 / 58.0 / 22.7 A&E: 56.0 / 58.0 / 22.4 Professional care: GP: 71.3 / 71.0 / 17.0 A&E: 70.0 / 71.0 / 17.8

	Health status: 258 patients completed the health status questionnaire (GP patients = 163, A&E patients = 95) **GP Number of patients (%)** / **A&E Number of patients (%)** Cured: 88 (54) / 47 (49) Improved: 48 (29) / 36 (38) Same: 22 (13) / 10 (11) Worse: 5 (3) / 2 (2) Had reattended A&E department for treatment of same complaint: 19 (12) / 9 (9) Had reattended own GP for treatment of same complaint: 40 (25) / 21 (22) Had original diagnosis subsequently changed: 4 (2) / 2 (2)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Socioeconomic class by General Medical Services (which are provided to the poor) 3 Yes, hospital's mainframe computer, patient satisfaction questionnaire, health status questionnaire 4 Within 30 days of first visit for reattendance, after 30 days follow up health status questionnaire 5 Yes, by time of arrival and status of 'semiurgent' or 'delay acceptable' 6 St. James's Hospital, Dublin, Ireland
Commentary	No data on consultation length of time. Low response rate for health status questionnaire.
Research implications	This study has shown that by comparison with the usual Accident and Emergency staff general practitioners investigate fewer patients, refer to other hospital services less often, more frequently refer patients back to their own general practitioner for follow-up, admit fewer patients, and prescribe more often. They do so with no apparent effect on patient outcome or on their subsequent use of hospital services. The study provides no explanations for these differences, which will be the subject of further research. Reasons for the more efficient performance of general practitioner staff might include their additional years of experience, their training in general practice, or their greater familiarity with community services. Indeed, the higher prescribing rates by the general practitioners may represent a different approach to the management of non-emergency patients, which itself warrants further exploration.

ID, origin, authors (year)	608, UK, Jackson, T.L. and Beun, L. (2000)
Aims	To study prospectively the outcome of conservative and surgical treatment of chalazia provided by medical and nursing staff *Workforce:* Senior nurse; primary care *Feature:* Substitution of senior house officer (SHO) for senior nurse *Outcome:* Success/complication rates of chalazion treatment, patient satisfaction and pain measure
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Prospective cohort study 2 All patients attending a district eye hospital for treatment of chalazion 3 129 patients, 217 visits; 170 visits where outcome could be determined 4 From initial visit through December 1995 5 Postal questionnaire or telephone call; recruitment began in January 1995 and data collection completed in December 1995.
Results Quantitative results	Patients who had surgical treatment with a nurse reported significantly lower amounts of pain and were more satisfied with explanation of treatment and treatment overall. *Outcome of treatment by doctor and nurse after first visit in these patients with known clinical outcomes* **Treatment** / **Patients treated exclusively by a nurse: success rate (patient number)** / **Patients treated exclusively by a SHO: success rate (patient number)** / ***p*-value** Conservative / 43% (28) / 13% (23) / 0.030 Surgical / 64% (28) / 83% (18) / 0.197 Overall success / 54% (56) / 44% (41) / 0.413 *Outcome of treatment after the first visit with I&C (surgical treatment) and conservative groups combined and including those patients lost to follow-up (patient numbers)* **Outcome of treatment** / **Nurse treatment** / **SHO treatment** / **Total** Success / 30 / 18 / 48 Lost to follow-up / 19 / 13 / 32 Failure / 26 / 23 / 49 Total / 75 / 54 / 129

	Patient satisfaction at most recent visit **Patient satisfaction** / **Patients treated by a nurse** / **Patients treated by an SHO** / ***p*-value** Mean pain score in cm (SD) / 1.7 (2.0) n=30 / 3.8 (2.7) n=23 / 0.003 *Q: How adequate was the explanation of your diagnosis? (%)* / *n=43* / *n=29* / 0.001 Very well explained / 58 / 24 Well explained / 40 / 52 Badly explained / 0 / 14 Not explained / 2 / 10 *Q: How adequate was the explanation of your treatment? (%)* / *n=43* / *n=29* / 0.003 Very well explained / 61 / 31 Well explained / 35 / 45 Badly explained / 5 / 14 Not explained / 0 / 10 *Q: Overall how do you rate the treatment you received at your last treatment? (%)* / *n=43* / *n=29* / 0.05 Very well explained / 81 / 52 Well explained / 16 / 35 Badly explained / 2 / 14 Not explained / 0 / 0
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 No difference in groups in age, sex or duration of symptoms; no patients known to be diabetic 3 Yes 4 Not beyond December 1995 5 None 6 One hospital in the UK
Commentary	No randomisation. Study was focused on treatment options and not on substitution; substitution was a secondary concern
Research implications	Study needs to be done with randomisation and with just one form of treatment

ID, origin, authors (year)	912, UK, Kinley, H., Czoski-Murray, C., George, S. *et al.* (2001)
Aims	To determine whether pre-operative assessment carried out by an appropriately trained nurse (ATN) is equivalent in quality to that carried out by pre-registration house officer (PRHO). To assess whether pre-assessments carried out by ATNs and PRHOs are equivalent in terms of cost. To determine whether assessment carried out by ATNs are acceptable to patients. To investigate the quality of communication between senior medical staff and ATNs. *Workforce:* Appropriately trained nurses (ATN); secondary care *Feature:* Substitution of pre-registration house officers (PRHO) for ATNs *Outcome:* Patient satisfaction
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Randomised controlled trial 2 All patients attending at one site for assessment prior to general anaesthetic for elective general, vascular, urological, or breast surgery were potentially included. Patients who were interviewed had to have signed and written consent form. 3 1907 patients were randomised, 1874 patients completed the study with full evaluation and 42 interviews were conducted. 4 Within 12-month time frame of study 5 12 months of data collection from interviewers and specialist registrars in anaesthetics
Results Quantitative results	For the RCT, looking at the history taking, examination, and test-ordering skills of ATNs and PRHOs, there was no difference found in the skills and both groups did too much or too little in each category. Only qualitative data were given for patient satisfaction.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 For surgery type, sex, and age 3 Yes for RCT; for patient satisfaction open-ended questions were given but not limited. 4 Not beyond initial visit and post-op interview. 5 Yes, for 42 interviews, 50% men, 50% women, 24 were from the two Southampton sites, 18 were from the Doncaster and Sheffield sites, 22 had seen an ATN and 20 had seen a PRHO. 6 2 sites in Southampton, 1 in Doncaster, and 1 in Sheffield. All four are NHS hospitals.
Commentary	Only 3 ATNs were involved in the study versus 87 PRHOs; this means that the results could be from the fact that stronger more confident nurses would be likely to seek extra training and sign up for this study. The study needs to be done in more areas and with more ATNs. Too few patients at Doncaster site, which may have made ATNs look less competent because there were not enough patients to examine. Patient satisfaction needs to be quantitatively measured instead of qualitatively to find true results of substitution.
Research implications	Quantitative study to measure patient satisfaction needs to be done. Substitution appeared successful, but a larger trial needs to be conducted using more ATNs.

ID, origin, authors (year)	914, Brazil, Lassner, K.J., Chen, C.H.C., Kropsch, L.A. *et al.* (1995)
Aims	To assess whether trained nursing personnel could provide IUD services as safely and effectively as physicians *Workforce:* Nurses with a university degree, technical nurses, auxiliary nurses or nurse trainees; primary care and family planning *Feature:* Substitution of trained nurse personnel to perform IUD insertions to low-income families *Outcome:* Insertion failures, insertion complications, pain at insertion, use effectiveness including rates of continuation and termination due to pregnancy, expulsion or removal, patient complaints
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised control trial 2 Inclusion criteria: all women who were eligible to receive the IUD and who voluntarily requested IUD insertion at the Centro de Pesquisas de Assistencia Integrada a Mulher e a Crianca (CPAIMC) clinic Exclusion criteria: CPAIMC's standard contraindications to IUD insertion: suspected or confirmed pregnancy, uterine abnormality, abnormal cervical cytology, pelvic infection but no history of pelvic inflammatory disease (PID) 3 1711 women requesting IUD insertion; 860 inserted by physicians and 851 inserted by nurses 4 November 1984 to 30 June 1987 5 Data recorded at follow-up visits, gynaecological examination results
Results Quantitative results	*Percentages of women with failure/complication/pain at insertion*

Type of provider	**Insertion failure**	**Complication at insertion**	**Pain at insertion**
Total	2.3	1.8	9.0
Physician (n=860)	1.3	1.7	10.8
Nurse (n=851)	3.3	1.8	7.1
	p = 0.005	p = 0.977	p = 0.008

Rate of insertion failure by type of provider

Characteristics of women	**Physician**		**Nurse**		**Total**	
	%	***n***	***%***	***n***	***%***	***n***
Parity						
0	3.4	117	11.6	146	8.0	263
Greater than or equal to 1	0.9	743	1.6	705	1.2	1448
	p = 0.027		p < 0.001		p = .005	
Women's age groups						
15–24	0.4	279	2.6	273	1.5	552
25–29	1.0	313	3.9	281	2.4	594
30–48	2.6	268	3.4	297	3.0	565
	p = 0.052		p = 0.067		p = 0.215	
Years of education						
0–4	1.5	328	1.4	294	1.5	622
5–8	1.4	342	3.1	322	2.3	664
>8	0.5	190	6.0	235	3.5	425
	p = 0.578		p = 0.013		p = 0.603	
Total	1.3	860	3.3	851	2.3	1711

Rate of termination by type of provider

Continuation and type of termination	Physician %	(SE)	Nurse %	(SE)	Total %	(SE)
Continuation rate	74.4	(2.0)	75.2	(1.9)	74.9	(1.4)
Termination rate	25.6	(2.0)	24.8	(1.9)	25.1	(1.4)
Pregnancy (cumulative)	1.4	(0.5)	1.0	(0.5)	1.2	(0.3)
Expulsion (cumulative)	5.3	(0.9)	5.0	(0.8)	5.1	(0.6)
Removal* (cumulative)	21.6	(2.0)	20.8	(1.9)	21.1	(1.4)
Reasons for removal:						
• medical reasons	8.7	(1.4)	6.2	(1.1)	7.4	(0.9)
• planned pregnancy	6.3	(1.2)	5.8	(1.1)	6.1	(0.8)
• other reasons	8.2	(1.4)	10.2	(1.5)	9.2	(1.0)

* Total rate of termination is less than the sum of the 3 rates of reasons for termination because these 3 rates were computed as gross rates without considering competing risks. Likewise, the 3 rates of reasons for removal were computed as gross rates.

Percentage having a complaint

Type of provider	Before IUD acceptance (1)	After IUD acceptance (2)	Difference (2 – 1)
Physician (780)	20.5	46.3	25.8
Nurse (771)	17.1	42.2	25.1
Total	18.8	44.3	25.5

This study concludes that trained nurses using the standard training and operating procedures at CPAIMC provided IUD services to family planning clients as safely and effectively as physicians. IUD use-effectiveness and side effects appear unrelated to whether the devices were inserted by physicians or nurses. However, if a nulliparous woman requests an insertion, to minimise risk of insertion failure that insertion should be performed by a physician or more experienced nurse with close medical supervision.

Quality appraisal

1 Case mix adjustment
2 Other adjustment
3 Uniform data collection
4 Participant follow-up
5 Random sampling
6 Geographical dispersal

1 N/A
2 N/A
3 Yes
4 Loss to follow-up: 121 (7.2%) did not return to clinic. 46% of all cases (47.2% of physicians' clients and 44.7% of nurses' clients) were censored at less than one year because they were lost to follow-up.
5 Yes, by one of the 11 physicians or 13 nurses at the clinic
6 One central clinic for research on integrated maternal and child care in Rio de Janeiro

Commentary	A small number of insertions included in the study were performed by nurse trainees from other institutions. All nurses received different amounts of training. In spite of the clients being randomly allocated by type of providers, some socio-demographic characteristics and medical conditions of the two groups were significantly dissimilar. Clients of physicians had higher mean parity and fewer years of schooling, higher histories of PID and sexually transmitted disease (STD) than nurses' clients. One possible explanation for these differences relates to differences in the recorded rates of PID and STD among the clients of nurses and physicians. If nurses were relatively poor at accurately detecting PID and STD then they could have been more likely to misdiagnose symptoms related to STD and PID, and so exclude the misdiagnosed women from the study. These phantom PID exclusions could have caused the nurses' clients to have significantly lower average parity and significantly higher average education than the physicians' clients since groups with higher parity and lower education tend to have higher rates of STD. Because the life-table analysis excluded the 7.2% of women who never returned to clinic after the initial visit but included the 46% of women who were censored at less than one year because they were lost to follow-up, the results of that analysis may not reflect the true levels of IUD use-effectiveness.
Research implications	Repeat the trial using: intervention groups of university-trained nurses and/or technical nurses in groups on their own, blinded assessment of participants in the trial, concealed randomised procedure by investigator not involved in the trial.

ID, origin, authors (year): 409, UK, Mann, A.H., Blizard, R., Murray, J. *et al.* (1998)

Aims

To evaluate the extended role for practice nurses in improving the outcome of depression through two specially designed interviews running in parallel
Workforce: Practice nurses; primary care
Feature: Extending role
Outcome: Change in health status; change in Beck Depression Inventory (BDI) and DSM-III criteria for major depression

Methods

1 Design
2 In-/exclusion
3 Sample size
4 Follow-up time
5 Data collection: source and period

1 Naturalistic random allocation study
2 GPs to refer those whom they thought were depressed; age 18–74 years who have been depressed for at least 4 weeks; those currently receiving treatment from their GP for depression or presenting with a new episode were included. Patients were excluded if they had suicidal ideation, with depression representing a phase in a manic-depressive psychosis, and those currently receiving treatment for depression from specialist psychiatric services.
3 577 patients in total, 524 patients completed through follow-up
4 Follow up at 4 months
5 Last 4-month follow-up done in August 1993. BDI and DSM-III used for measurement.

Results

Quantitative results

There was no difference between the groups, proving the hypothesis incorrect.

Changes in BDI score over 4 months, comparing intervention with control groups in Study 1

	Number		Entry		Outcome	
	Entered	***Complete***	***DSM-III (%)***	***BDI mean***	***DSM-III (%)***	***BDI mean***
Control	82	74	78	18.47	24	11.53
Intervention	74	65	86	18.62	27	11.52

Changes in BDI score over 4 months, comparing intervention with control groups in Study 2

	Number		Entry		Outcome	
	Entered	***Complete***	***DSM-III (%)***	***BDI mean***	***DSM-III (%)***	***BDI mean***
Control	148	134	86	20.75	27	10.15
Intervention	271	251	80	21.14	31	10.87

Quality appraisal

1 Case mix adjustment
2 Other adjustment
3 Uniform data collection
4 Participant follow-up
5 Random sampling
6 Geographical dispersal

1 and 2 No, both groups were naturally similar
3 Yes, done by nurses at appointments
4 Not beyond 4-month follow-up
5 Yes, patient files numbered consecutively and were randomly allocated to groups by random number tables.
6 20 general practices distributed throughout England

Commentary

There was a high prescription rate of antidepressants which could have affected both groups. The study should be conducted by separating patients taking antidepressants from those who are not; or by not allowing health care workers to prescribe antidepressants to fully measure the effect of the added nurse counselling.

Research implications

Could repeat study to explore patient satisfaction to determine if the either group does a better job.

ID, origin, authors (year)	50, England, Moore, S., Corner, J., Haviland, J. *et al.* (2002)
Aims	To assess the effectiveness of nurse-led follow-up in the management of patients with lung cancer *Workforce:* Clinical nurse specialists; secondary care *Feature:* Nurse substitution in doctor's role of outpatient follow-up *Outcome:* Quality of life and patient outcome (secondary outcomes of mortality/survival rates, symptom-free survival, and progression-free survival)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Randomised controlled trial 2 Patients with lung cancer who completed initial treatment and are expected to survive 3 or more months 3 203 patients consented and qualified for randomisation 4 12 months 5 Standardised questionnaires produced by the European Organization for Research and Treatment of Cancers (EORTC), no dates of data collection provided
Results Quantitative results	*Patient satisfaction*

Item	**Nurse-led**	**Conventional**	***p*-value**
3 months	*(n=75)*	*(n=71)*	
Organisation of care	81.3 (75.0–93.8)	71.9 (65.6–78.1	<0.001
Information and advice	77.1 (69.8–89.6)	68.8 (58.3–75.0)	<0.001
Personal experience of care	77.3 (75.0–95.5)	75.0 (68.2–80.1)	0.002
Satisfaction with care	78.4 (61.6–100)	70.0 (51.1–79.5)	0.005
How would you rate your support overall?	93.0 (80.0–100)	78.0 (57.0–94.0)	0.002
6 months	*(n=52)*	*(n=55)*	
Organisation of care	83.3 (75.0–93.8)	75.0 (68.8–78.1)	<0.001
Information and advice	75.0 (67.9–85.4)	66.7 (58.0–75.0)	<0.001
Personal experience of care	79.5 (72.7–97.7)	75.0 (68.2–77.3)	0.001
Satisfaction with care	79.5 (65.9–98.3)	75.0 (58.7–89.2)	0.11
How would you rate your support overall?	89.0 (82.8–98.3)	83.0 (64.5–96.3)	0.04
12 months	*(n=27)*	*(n=29)*	
Organisation of care	81.3 (75.0–96.9)	75.0 (70.3–83.3)	0.01
Information and advice	75.0 (70.8–91.7)	68.8 (64.6–77.1)	0.01
Personal experience of care	79.5 (75.0–100)	75.0 (70.2–87.5)	0.03
Satisfaction with care	82.5 (72.7–100)	76.1 (64.2–85.8)	0.13
How would you rate your support overall?	93.0 (77.0–98.0)	81.5 (70.0–95.0)	0.08

At 3 months patients were significantly more satisfied with nurse led care. At 3 months 53/75 (78%) patients randomised to nurse follow up said they would prefer nurse led care if asked to choose, but only 11/71 (17%) of patients who received conventional medical follow up would prefer to see a doctor only.
Although no evidence showed a difference in objective progression, evidence showed that the nurses recorded symptomatic progression sooner than the doctors.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Stratified at randomisation for type of lung cancer, stage of cancer, comorbiditites, hospital, treatment intent, etc. 3 Yes, EORTC questionnaire given at same time intervals 4 Yes at 3, 6, and 12 months (203 patients at start, 156 at 3 months, 113 at 6 months, 60 at 12 months) 5 Yes, an independent trials office was responsible for randomisation and stratification 6 South-eastern England
Commentary	Difficult to truly compare competency, knowledge base and efficiency of nurses because doctors followed up with patients every 2–3 months, while nurses followed up at least once a month. There was no mention of patient attitudes to the title of 'doctor' or 'nurse' which could affect the scores given to the doctors and nurses (i.e. gender beliefs and division of labour, general dissatisfaction and frustration with doctors). The rate of attrition was high because of death or disability. Such difficulties with recruitment and attrition are recognised problems of research studies conducted with very ill and dying patients. The number of outcomes measured in this study would imply that some findings may have occurred by chance.
Research implications	Can nurses provide the same level of care if patients are seen at the same interval as doctor visits? Do these findings hold true for other specialty care? Do nurses want this increased responsibility? Are doctors comfortable in handing over this responsibility? Does this study simply show that doctors should be doing more?

ID, origin, authors (year)	220, USA, Pinkerton, J. and Bush, H.A. (2000)
Aims	To compare perceived health and satisfaction with care in a managed care system in two groups of patients *Workforce:* Nurse practitioners (NP); primary care *Feature:* Substitution of physicians by NP in a managed care setting *Outcome:* Perceived health and patient satisfaction
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Cross-sectional 2 Ambulatory patients attending an outpatient clinic, able to read and understand English and were cared for by NP or a physician, but not both, and who presented with a diagnosis of diabetes mellitus and/or hypertension were included in the study. 3 160 clinic patients (80 in each group) aged 18 to 89 in a managed care setting 4 Not stated 5 SF-20 Health Survey and the NP Satisfaction Instrument (NPSI) and demographic data sheet
Results Quantitative results	The SF-20 total score means (values not quoted) for NPs and physician's groups tested with the *t*-test for dependent samples resulted in no significant difference (t = –0.95, df = 148, p = 0.34), the inference being that the perception of health for both groups was the same. The NPSI scores (values not quoted) were tested using the *t*-test for independent samples. Results indicated no significant difference in the NP's and physician's groups (t = –0.92, df = 149, p = 0.60), implying that patient satisfaction with care was the same for both groups. The findings may mean that NPs placed in managed care environments can be expected to perform as effectively as they have in non-managed care environments. The findings also imply that NPs could be placed in managed care settings where there is no availability of primary care physicians, not only in the interest of cost containment of health care services, but of actual health care and patient satisfaction, and that the managed care system may prove to be the catalyst to deeper changes in the practice of both nurses and physicians. As knowledge and skills continue to expand, each discipline will change, with further shifting and sharing of role components.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 and 2 None made 3 Yes 4 Loss of follow up: 14 on the SF-20. 5 No, availability sampling used 6 One public hospital in a large south-western city in the USA
Commentary	Response bias may have occurred since 40 of the patients who were seen by the physician and 12 who received care from the NP preferred to be seen by just the physician, while 88 participants reported that they would be seen by either the NP or the primary care physician. These patients may have tried to present a more favourable side to the investigator regarding the provider or even the outpatient clinic. However, if this occurred, it occurred for both groups.
Research implications	A qualitative study should be undertaken to understand why patients report a preference for care administered by either the NP or the physician. Additional studies comparing NPs and physicians in a managed care setting are needed before a conclusion could be formulated. This study did not investigate the factor of cost containment and further research into the economics of health and nursing care should be undertaken.

ID, origin, authors (year)	212, USA, Pioro, M.H., Landefeld, S.C., Brennan, P.F. *et al.* (2001)
Aims	To compare care delivered by nurse practitioners and house staff and to investigate whether nurse practitioners can admit and manage general medical patients *Workforce:* Nurse practitioners (NP); secondary care *Feature:* Substitution of medical housestaff (MH) by NPs and a medical director in general medical wards *Outcome:* Length of stay, hospital charges, costs, number of consultations to other services, adverse events: transfers to intensive care, in-hospital and 30-day mortality, hospital-acquired complications, patient assessments of care, changes in activities in daily living (ADL), health status (SF-36), symptom severity, and patient assessment of care.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Randomised controlled trial (RCT) 2 Patients aged 18–69 years admitted through outpatient facilities or the emergency room were eligible. Excluded patients were: those admitted to or from intensive care units or other specialty units (telemetry ward, coronary step-down unit, haematology–oncology ward, bone marrow transplant ward, HIV ward). Patients admitted during 'off-hours' (between 17:00 and 07:30 weekdays and throughout the weekend) were also initially excluded. However, beginning July 1994 patients admitted from 17:00 to 07:30 Monday to Friday were also randomised. 3 381 general medical patients, 193 of which were assigned to the NP-based care and 188 to housestaff care 4 March 1994 to September 1995 5 Medical records, patient interview and hospital administrative database, National Death Index registry for all deaths recorded in the USA until 31 December 1995

Results
Quantitative results

Primary outcomes using intention to treat (ITT) analysis (difference between NP-based care and housestaff care were not significant (p >0.10), either by ITT or actual treatment analysis)

	NP-based care	Housestaff care	NP-housestaff (95% CI)
Length of stay (mean) days	5.0	5.3	–0.3 (–1.2, 0.6)
Total hospital charges (mean) US$	8854	9426	–572 (–2704, 1560)
Total ancillary charges (mean) US$	4960	5358	–399 (–1820, 1023)
Cost (mean) US$:			
• Pharmacy	393	388	5 (–161, 172)
• Radiology	382	460	–78 (–216, 62)
• Laboratory	640	639	1 (–205, 208)
• Respiratory therapy	105	150	–45 (–115, 24)
No of consultations/patients (mean)	1.4	1.4	0.0 (–0.2, 0.3)
Transfer to intensive care unit (%)	3.6	6.9	–3.3 (–7.8, 1.2)
In-hospital mortality (%)	1.6	1.1	0.5 (–1.8, 2.8)
30 days post-discharge mortality (%)	3.6	3.2	0.4 (–3.2, 4.0)
> 1 hospital-acquired complication (%)	5.3	8.6	–3.3 (–8.4, 1.8)
Overall adverse event rate, %	7.5	11.8	–4.3 (–10.2, 1.6)
Discharge disposition (%):			
• Home	92.6	96.2	–3.7 (–8.2, 1.0)
• Skilled nursing facility	6.4	3.2	3.1 (–1.1, 7.5)
• Left against medical advice	1.1	0.5	0.5 (–1.2, 2.4)

	NP-based care	Housestaff care	NP–housestaff (95% CI)
Outcomes at discharge	*n=106*	*n=115*	
Improved from admission in no. of dependent ADL (mean)	0.3	0.2	0.1 (–0.2, 0.4)
Improved from admission in no. of dependent IADL (mean)	1.0	1.2	–0.2 (–0.8, 0.4)
Decrease from admission in symptom severity (mean)	5.5	5.1	0.4 (–0.5, 0.3)

Outcomes at 6 weeks post-discharge	*n=76*	*n=86*		
Improved from admission in no. of dependent ADL (mean)	0.1	0.2	–0.1	(–0.5, 0.3)
Improved from admission in no. of dependent IADL (mean)	1.4	2.1	–0.7	(–1.4, 0.1)
Decrease from admission in symptom severity (0–10) (mean)	4.0	4.7	–0.7	(–2.6, 1.2)
Patient assessment of care				
Overall rating (0–100) (mean)	84.7	80.7	4.0	(–3.0, 11.0)
Patient perceived problems (0–100) (%)	10.1	10.1	0.0	(–9.3, 9.3)
Physician and nursing care (0–100) (mean)	77.3	76.7	0.6	(–6.1, 7.3)
Mean improvement from admission in SF-36 scores				
Single item health status	4.7	2.9	1.8	(–3.0,12.6)
Physical functioning	–3.3	–0.8	–2.5	(–10.0, 5.0)
Social functioning	4.0	1.1	2.9	(–5.8, 11.6)
Role functioning (physical problems)	6.1	3.1	3.0	(–9.6, 15.6)
Role functioning (emotional problems)	4.4	3.7	0.7,	(–13.8, 15.2)
Mental health	3.4	3.5	–0.1	(–6.0, 5.8)
Vitality	4.8	5.1	–0.3	(–7.3, 6.7)
Pain	15.0	12.1	2.9	(–6.4, 12.2)
General health	1.4	–2.4	3.8	(–1.9, 9.5)

NP-based care can be implemented successfully in teaching hospitals and compared to housestaff care, may be associated with similar costs and clinical and functional outcomes. However, there may be important obstacles to increasing the number of patients cared for by NPs, including physician concerns about NPs capabilities and NPs' limited flexibility in managing varying numbers of patients and accepting off-hours admissions. Thus while it is unlikely that NPs can replace housestaff, the findings indicate that NP-based care can complement house-staff care and reduce the number of housestaff needed while providing similar levels of service.

Quality appraisal	
1 Case mix adjustment	1 N/A
2 Other adjustment	2 N/A
3 Uniform data collection	3 Yes
4 Participant follow-up	4 Medical record reviews were complete for 374 (98%) of patients. Records were unavailable for 7 (2%) patients. Interview data were complete for 69% of patients at admission, 58% at discharge and 43% 6 weeks after discharge.
5 Random sampling	5 Yes
6 Geographical dispersal	6 One university hospital in Cleveland

Commentary	The high rate of patient crossover, (90 of the 193 patients (47%) assigned to the NP ward were admitted to housestaff wards and 1 patient assigned to housestaff care was admitted to NP ward) due to attending physicians and NP requests and lack of beds on the NP ward may have introduced selection bias. Doctors wanted the flexibility to pre-empt randomisation because of concerns that certain patients may be 'too sick' to be managed by PNs and that admitting patients to the NP ward might increase their involvement in 'off hours' management. The power to detect clinically meaningful differences between the two groups was relatively low for some endpoints. For example, while the power to detect a 30% difference in length of stay, charges or costs was roughly 80% the power to detect a 50% difference in rats of adverse events was only 20%. Because of the nature of the intervention patients and interviewers were not blinded to treatment assignments. The assessment of the cost of care principally reflected the use of discretionary hospital resources (e.g. bed days, diagnostic tests) and did not explicitly consider differences in NP and housestaff salaries, the costs of the medical director, nor the costs of providing off-hours coverage by residents.
Research implications	The generalisability of the findings to other teaching hospitals should be established. Several important organisational issues need to be considered in implementing NP-based care, including doctors' perceptions of NP's capabilities, especially among patients perceived as being 'very ill' and the decreased flexibility of NPs to accommodate off-hours admissions and wide fluctuations in numbers of patients.

ID, origin, authors (year)	155, UK, Pritchard, A., Kendrick, D. (2001)
Aims	To evaluate practice nurse and health visitor management of patients with acute minor illnesses, monitor the effect on general practitioner workload, and describe the range of conditions seen by nurses *Workforce:* Practice nurses (PN) and health visitors (HV); primary care *Feature:* Substitution of GPs for PNs and HVs *Outcome:* Patient satisfaction, re-admission for same problem within 2 weeks
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Before and after 2 Patients with 'urgent' appointments and were being seen for: colds, influenza, coughs, asthma, sore throats, earache, high temperatures, diarrhoea, vomiting, cystitis, sore eyes, conjunctivitis, skin rashes, infections, bites, stings, cuts, bruises, ingrowing toenails, thrush, shingles, allergies, hay fever, nose bleeds, emergency contraception, and mouth ulcers. 3 1900 patients 4 None 5 GP-led appointment data from January to June 1998; NP- and HV-led appointment data from January to August 1999; information included date of appointment, diagnosis, age, who first saw the patient, referrals to GPs and urgent re-consultations for the same illness within 2 weeks, etc. Data from 1998 were collected from appointment sheets and supplemented by computer and paper records. 1999 data were collected similarly, the data were then downloaded from the patient record system and imported to an Access database. Two patient questionnaires: the Consultation Satisfaction Questionnaire (CSQ) and the Patient Enablement Index (PEI).
Results Quantitative results	1999 re-consultation rate (7.4%) was lower than that in 1998 (9.2%). Patient satisfaction survey: in general, higher satisfaction was found with the HV, and there was no difference in the GP and PN. *Satisfaction score:* GP (n=227,response rate=72.5%); PN (n=140,response rate=72.9%); HV (n=22,response rate=88.0%); Total (n=389,response rate=74.4%) **GP** **PN** **HV** **Total** CSQ 73.3 72.4 77.7 73.3 PEI 25.0 12.5 50.0 25.0
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 and 2 No adjustments made 3 Yes 4 Not beyond 2 weeks of visit 5 No, used patients from practice's patient list 6 Nottingham, England
Commentary	Patients expressed high levels of satisfaction with the acute minor illness service. Many remarked that the opportunity to see a nurse was especially useful if it reduced the waiting time for an appointment. Patients were equally satisfied with being seen by PNs or GPs, with HV appointments scoring higher than those with both GP and PN. Nurses have different but complementary skills and approach to doctors, and have a unique contribution to make to the primary health care team. Research may not be generalisable; only done at one practice.
Research implications	Studies on different approaches to care between doctors and nurses that cause satisfaction rates.

ID, origin, authors (year)	253, England, Reynolds, H., Wilson-Barnett, J., Richardson, G. (2000)
Aims	To investigate differences between care provided by the hospital-based Parkinson's disease nurse specialist compared with the consultant neurologist *Workforce:* Parkinson's disease nurse specialists (PDNS) *Feature:* Substitution of PDNSs for consultant neurologists *Outcome:* Quality of life, patient satisfaction
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Seen by consultant at least once for medical assessment and confirmation of diagnosis of idiopathic Parkinson's disease (PD), not previously seen by a PDNS, able to understand requirements of study and give informed consent, no clinical evidence of dementia, new referral to clinic 3 108 patients in 3 outpatient centres 4 12 months 5 Patient survey done at visits over 12 months, dates of collection not stated. Survey included Parkinson's disease questionnaire, hospital anxiety and depression scale, SF-36 health status questionnaire, functional disability questionnaire, and patient satisfaction survey.
Results Quantitative results	Only 2 out of 22 dimensions reached statistical significance (p = 0.05) when analysis of differences was performed for physical disfunctioning (p = 0.02) and general health (p = 0.02), both measured by SF-36 and both favoured the consultant only group. All groups maintained their baseline HAD scale results except for the group referred to PDNS which increased to mild median anxiety. Median social disability scores decreased in all groups except where patients were referred to PDNS where median social disability improved at the end of the study. Median physical activity improved slightly in the PDNS-only group and remained the same in the consultant-only group. Median physical activity deteriorated in the other two groups during the study. Median self-care improved considerably in the group where patients were referred to PDNS, self-care remained the same in the PDNS-only group, and deteriorated in the other two groups. No significant differences were shown on the patient satisfaction survey between any of the groups.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 and 2 Randomised based upon age, time since diagnosis, sex, living alone or with another person, etc. 3 Consistent tool used for data collection, but number of times data was collected per patient was not consistent. 4 Yes, over a 12 month period 5 Yes 6 3 outpatient centres in England
Commentary	Combined care was provided for those with multiple problems. Outcomes may reflect the deterioration in the condition more than the benefit from interventions. Sicker patients required more interventions and more support. Compared with other groups, patients receiving care from both consultant and nurse showed deterioration in some median health outcome scores despite receiving more interventions from both specialists. Consultant and PDNS consultations covered similar information but varied in focus and time spent with patient. Potential bias arose in selection of study sites in that PDNS promoted clinics to the study that were perceived as particularly 'good'. Lack of a control clinic with no PDNS available. Sample size decreased throughout study due to disability and deteriorating health of patients. Complementing rather than substitution was seen as the way for the future, though?
Research implications	Are PDNSs needed? Can PDNSs takeover the workload of consultant neurologists working with PD patients? Is "team" care more efficient for patients? Is 'team' care better at reducing patient dissatisfaction and morbidity rates?

ID, origin, authors (year)	540, USA, Rifkin, W.D., Conner, D., Silver, A. *et al.* (2002)
Aims	To compare medical care provided by hospitalists and primary care physicians to patients with community-acquired pneumonia in order to identify specific practices that might explain the improved efficiency of care provided by hospitalists *Workforce:* Hospitalists, primary care physicians; tertiary care centre *Feature:* Substitution/specialisation (hospitalists – cared for patients only during hospitalisation; primary care physicians – also provided care after discharge) *Outcome:* LOS, re-admission rates, mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective chart review 2 Inclusion: Adults admitted to the study centre and coded on discharge with a principal diagnosis of community-acquired pneumonia. Exclusion: Patients with known HIV, lung cancer, active mycobacterium tuberculosis, or prior hospitalisation within 7 days and those who required mechanical ventilation during hospitalisation or had a length of stay longer than 14 days. 3 9 hospitalists, 56 physicians, 455 patients 4 In-hospital 5 Three RNs reviewed the medical records for patients admitted from 1 January 1998 to 1 January 1 1999. Also reviewed the dictated radiology reports for each patient's initial chest X-ray film. Administrative databases were used to collect the outcome data.
Results Quantitative results	Patients cared for by hospitalists had a shorter LOS compared with those cared for by primary care physicians. The mean crude LOS was 5.9 (median = 5) for hospitalist patients and 7.0 days (median = 6) for primary care physician patients. An adjusted mean LOS was 5.6 days for hospitalist patients and 6.5 days for primary care physician patients (p = 0.001). Unadjusted hospital re-admission rates at 15 and 30 days were higher for hospitalist patients but were not statistically significant. Mortality was higher for patients of primary care physicians – with an adjusted odds ratio for hospitalist inpatient mortality of 0.37 (ratio not statistically significant).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Multivariate adjustments accounting for insurance status (self-pay, Medicaid, private or Medicare), age (≤66–85, >85), Pneumonia Severity Index status (low, medium, high risk), whether the patient died and whether the patient came from a skilled nursing facility 3 Yes 4 Charts for 3 patients were unavailable; 73 patients excluded because they were transferred from a sub-acute care centre with nosocomial pneumonia; 4 patients had HIV; 3 patients with post- obstructive pneumonia from lung cancer; 16 hospitalised within previous 7 days and 4 were undergoing mechanical ventilation. 5 No 6 One centre in New York
Commentary	Primary care physicians cared for patients who were on average 5 years older (p = 0.002) and more likely to have severity risk class 5 pneumonia (p = 0.02). It appears that an earlier switch can explain most of the efficiencies seen in the study from intravenous to oral antibiotics. Possibly this early conversion was facilitated by the fact that hospitalists, not primary care physicians, were on site. The process of care measures examined showed that hospitalists and primary care physicians practice similarly. Only looked at in-hospital events as outcomes. The index used for severity may be better for adjusting mortality than LOS. Hospitalists were twice as likely to discharge a patient with an abnormal measure of stability. Hence hospitalists are discharging patients more quickly but they are sicker.
Research implications	Further investigation of this relationship using longitudinal data from multi-centres. Follow-up of patients after discharge to examine the effects of discharging them more quickly but when they are sicker. Would the experience or training of the physician or hospitalist affect the results?

ID, origin, authors (year)	265, Canada, Rubin, S., Weins, L., Fingler, I. *et al.* (1996)				
Aims	To assess whether there was a difference in patient outcomes when femoral venous and arterial sheaths were removed post percutaneous transluminal cornonary angioplasty by medical doctors as compared to registered nurses *Workforce:* Registered nurses (RN) 75% diploma prepared, 25% baccalaureate degree in nursing; tertiary care *Feature:* Substitution of medical doctor (MD) by RNs to remove femoral venous and arterial sheaths post percutaneous transluminal coronary angioplasty (PTCA). *Outcome:* Complication rate: bleeding, haematoma formation, vagal reaction, use of analgesics and anxiolytics. The impact on nursing practice of nurses assuming this new task was examined but not in relation to patient outcomes.				
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Observational study 2 Patients admitted to a 12-bed Cardiology Interventional Unit (CIU) for PTCA 3 139 patients whose femoral sheaths were removed by MDs and 122 patients whose femoral sheaths were removed by RNs. 4 January 1993 to May 1993 5 Patient's charts held by the Health Records Department				
Results Quantitative results	**Outcome**	**Result**	**Chi square**	**df**	***p*-value**
	Bleeding	Bleeding was significantly greater in patients whose sheaths were removed by MDs.	10.51	1	<0.01
		14% patients bled from the groin when RNs were removing sheaths whereas 33% patients experienced bleeding when MDs were removing sheaths. The significant difference in bleeding occurred with the sheaths in situ, whereas no statistically significant difference in bleeding occurred with the clamp on and post-clamp removal.	3.93	1	<0.05
	Haematomas	No significant difference between nurses and MDs removing sheaths.			
	Vagal reaction	No significant difference between nurses and MDs removing sheaths.			
	Analgesic medication	When RNs removed the sheaths 74.4% patients received pre-sheath analgesic compared to 47.5% of patients when MDs removed the sheaths.	17.98	1	<0.01
	Anxiolytic medication	RNs gave significantly more patients anxiolytic medication pre-sheath removal than MDs.	32.18	1	<0.01
		27% patients given anxiolytic medication while the clamp was on had an incident of groin bleeding, whereas only 7.1% of patients who did not receive anxiolytic medication had an incident of bleeding.	9.98	1	<0.002
		More bleeding occurred post-clamp removal if an anxiolytic medication was given while the clamp was on. 29% of patients had bleeding post-clamp removal when given an anxiolytic while the clamp was on compared to 10% of patients who bled while the clamp was on and had not received an anxiolytic medication.	7.47	1	<0.006

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 N/A 2 N/A 3 Yes 4 Complete 5 No, convenience sampling used 6 One teaching hospital in Vancouver
Commentary	Differences in bleeding may be explained by the differences in patient populations. RN removed sheaths only from elective PTCA whilst MDs removed sheaths from both emergency and elective PTCA patients. Emergency patients may have received thrombolytic therapy making them more prone to bleeding after the procedure. Sheaths more often remained in situ overnight when MDs were removing them. This lengthy period of bed rest with the sheaths in situ often led to discomfort and restlessness due to long periods in the supine position and a longer time on anticoagulant therapy.
Research implications	Further research is needed to evaluate patient satisfaction with nurses' performance of sheath removal and patient perceptions regarding preparation for the procedure and pain and anxiety management at the time of sheath removal. Nurses assuming functions that were previously considered in the domain of medical practice can be a source of dissatisfaction and warrants ongoing discussion.

ID, origin, authors (year)	675, UK, Sakr, M., Kendall, R., Angus, J. *et al.* (2003)
Aims	To compare the clinical effectiveness and costs of minor injury services provided by nurse practitioners with minor injury care provided by an accident and emergency department *Workforce:* Nurse practitioners; primary care *Feature:* Substitution *Outcome:* Number of errors in clinical assessment, treatment, and disposal
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Before-and-after cohort study 2 Patients attending the A&E unit between 7 August 1996 and 19 November 1996 with a minor injury; patients attending the MIU between 1 September 1997 and 31 January 1998; patients who presented as a 999 call even with a minor injury were excluded. 3 1447 patients in A&E group; 1315 patients in MIU group 4 None 5 Double assessments were done for patients in the A&E unit between 7 August 1996 and 19 November 1996 and patients attending the MIU between 1 September 1997 and 31 January 1998. Researchers compared the 'research assessment', done by investigators or radiologists and considered the gold standard, to the clinical assessment.
Results Quantitative results	A NP minor injury unit can provide a safe and effective service for the treatment of minor injury. *Numbers (%) of patients with clinically significant errors* (in some patients errors may have been made in more than one category)

	A&E n=1447		**MIU** n=1315		***p*-value**
Number of patients with at least one significant error	191	(13.2)	126	(9.6)	0.003
Errors in history of injury	2	(0.15)	1	(0.1)	0.9
Errors in past medical history	32	(2.2)	5	(0.4)	<0.0001
Errors in examination	28	(1.8)	43	(3.3)	<0.03
Errors in follow-up	48	(3.3)	42	(3.2)	0.9
Errors in treatment	60	(4)	42	(3.2)	0.2
Errors in radiological interpretation:					
• False negative	9	(0.6)	4	(0.3)	
• False positive	4	(0.3)	6	(0.4)	0.7

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 No, but groups were comparable in the age, sex, method of presentation, and triage category 4. 3 Yes 4 No 5 No 6 Sheffield, UK
Commentary	NPs were equal and sometimes better compared to A&E doctors; NPs had lower waiting times. No patients were followed up to see if errors affected health outcomes. NPs made more follow-up appointments than A&E doctors, but seen as appropriate because there was no senior advice available. Researchers were not blinded as to which cohort they were investigating; could lead to bias.
Research implications	Repeat study with randomisation and patient follow-up. Can NPs work within an A&E and provide results?

ID, origin, authors (year)	929, Canada, Spitzer, W.O. *et al.* (1990)
Aims	To assess the effects of substituting nurse practitioners for physicians in primary-care practice. *Workforce:* Nurse practitioners and physicians; primary care *Feature:* Substitution: a randomised conventional group in which the primary clinical care was from a family physicians, and a randomised nurse practitioner group, whose primary care was provided by a nurse practitioner *Outcome:* Patient satisfaction, physical status, emotion status, social function, death
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Families in which one of the family members either made contact with one of the practices in the prior 18-month period or (during later interviews) identified the doctor as the family physician 3 1598 families, containing 4325 members 4 1 year: 1 July 1972 to 30 June 1973 5 For preparation period (1 December 1970 to 1 May 1971): questionnaire For baseline period (1 May 1st 1971 to 1 July 1971): patient status survey For comparison period (May 1st, 1971 to July 1st, 1972): patient status survey For experimental(1st July, 1971 to 1st, May 1972) and follow-up period (July 1st, 1972 to June 30th, 1973): day-sheet journal Data collection period is Dec 1st, 1970 – May 1st, 1972
Results Quantitative results	A nurse practitioner can provide first-contact primary clinical care as safely and effectively, with as much satisfaction to patients as a family physician. The levels of physical status remained closely similar in the patients in the two groups. The index of emotional function and social function were 58.3%, 57.9%; 83.2%, 83.9 % respectively. During the experimental period, the difference in crude death rates was not clinically or statistically significant. In the follow-up survey, 97% of patients in the conventional and 96% in the nurse practitioner group were found to be satisfied with health services received during the experimental period.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Two groups were found to be highly similar on physical function, ability to carry out usual daily activities and freedom from bed stability. The base-line health status of the two groups of patients showed only minor differences that were not statistically significant (at an α level of 0.05). 2 N/A 3 Uniform 4 7 families refused their assignment: 2 from the conventional group preferred care by nurse practitioners, 3 from the nurse practitioner group opposed the new concept, and 3 others in the nurse practitioner group had had a member under care by a doctor for a long-term problem. 0.9% of families in the conventional and 0.7% in the nurse practitioner group left the practice because of dissatisfaction. 5 Because a caseload half that of a family physician's was considered manageable for a nurse practitioner, the eligible families were stratified by practice of origin, and randomly allocated in a ratio of 2:1. 6 A middle-class suburban town of 85,000 population in east of Hamilton.
Commentary	Did not take into account the characteristics of the health providers.
Research implications	It is important in planning of health care delivery for regions where family physicians are in short supply.

ID, origin, authors (year)	271, UK, Sturgess, R.P., O'Toole, P.A., McPhillips, J. *et al.* (1996)
Aims	To evaluate the success rate and complications of percutaneous endoscopic gastrostomy (PEG) insertion performed with an endoscopy nurse practitioner, rather than a second doctor, carrying out percutaneous gastric puncture *Workforce:* Nurse practitioner (NP); secondary care *Feature:* Substitution of doctor for NP *Outcome:* Complication rate
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective cohort study 2 Patients were unselected requiring routine PEG insertions 3 100 patients, 50 in each cohort 4 3 months after insertion 5 Case notes reviewed, GPs and nursing homes contacted to provide information on complications, 30-day mortality, and 3-month outcome of PEG placement if patient could not come for follow-up
Results Quantitative results	With appropriate training NPs can be equally successful in PEG placement. Results and complications of PEG placements **Nurse assisted** / **Doctor assisted** Successful PEG insertion: 50 / 49 Complications: • Immediate: 2 mucosal bleed, respiratory arrest / 2 acute stridor, lost puncture • Late: 1 infection / 3 infection Deaths: • 30 days: 4 / 6 • 3 months: 9 / 8
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and2 No, groups were 'roughly similar' 3 Yes 4 3 months after PEG placement follow-up was completed for 57 patients; other data were collected from GPs and/or nursing homes 5 No 6 One unit in one UK hospital
Commentary	Only one NP; needs to be repeated with more nurses to avoid volunteer bias. No case mix adjustment No randomisation
Research implications	Needs to be repeated on a larger scale; more nurses, more hospitals, and more cases. Is there a difference in patient satisfaction between nurses and doctors?

ID, origin, authors (year)	68, Netherlands, Tijhuis G.J., Zwinderman, A.H., Hazes, J.M.W. *et al.* (2002)
Aims	To compare the clinical effectivenss of care delivered by a clinical nurse specialist, inpatient team care and day patient team care in patients with rheumatoid arthritis who have increasing functional limitations *Workforce:* Clinical nurse specialist (CNS); secondary care *Feature:* Substitution of CNS for inpatient team care and day patient team care in patients with rheumatoid arthritis (RA) who have increasing functional limitations. Identification of patients who would benefit most from one of the three types of care. *Outcome:* Functional status, quality of life, health utility, disease activity, patient satisfaction.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Patients were included if they were diagnosed with RA as defined by the 1987 American College of Rheumatology criteria (6) and increasing difficulty in performing activities of daily living over the previous 6 weeks. Exclusion criteria were medical complications of RA requiring immediate hospitalisation and inability to reach the hospital before 10:00 a.m. 3 210 patients with RA 4 December 1996 to January 1999 5 Health Assessment Questionnaire (HQA); McMaster Toronto Arthritis Patient Preference Disability Questionnaire; (MACTAR), RAND-36 Item Health Survey (RAND 36); Rheumatoid Arthritis Quality of Life (RAQol) questionnaire; Health Utility Rating Scale; Disease Activity Score (DAS). Patient satisfaction on a visual analogue scale (VAS).
Results Quantitative results	*Clinical outcome data at baseline (absolute values) and 52 weeks* (change scores from baseline, means (95% confidence interval))

	Baseline	**Week 52**		**Baseline**	**Week 52**
HAQ (0–3)			**RAQol (0–30)**		
Nurse specialist patients (NSP)	1.17 (0.65)	0.17 (0.03, 0.30)*	Nurse specialist patients	13.8 (7)	1.7 (0.3, 3.1)*
Inpatients (IP)	1.49 (0.71)	0.19 (0.06, 0.32)*	Inpatients	17.0 (6)	1.4 (0.1, 2.8)*
Day patients (DP)	1.54 (0.76)	0.36 (0.23, 0.50)*	Day patients	18.3 (7)	3.1 (1.6, 4.5)†
MACTAR weighted			**Rating scale**		
Nurse specialist patients	48.4 (3.7)	–4.3 (–6.8, –1.8)*	Nurse specialist patients	60.9 (17)	–10.1 (–14.0, –6.1)†*
Inpatients	47.2 (3.6)	0.6 (–2.0, 3.1)	Inpatients	54.7 (17)	–4.6 (–8.5, –0.8)*
Day patients	47.4 (3.7)	–5.3 (–7.9, –2.6)*†	Day patients	54.5 (19)	–10.9 (–15.1, –6.7)†
RAND Physical summary scale (0–100)			**Disease activity score**		
Nurse specialist patients	38.0 (21)	–15.7 (–21.5, –9.9)†	Nurse specialist patients	5.32 (1.24)	1.3 (0.9, 1.6)†
Inpatients	29.6 (17)	–10.4 (–16.0, –4.8)†	Inpatients	5.72 (1.17)	0.9 (0.6, 1.2)†
Day patients	28.2 (20)	–15.7 (–21.5, –9.9)†	Day patients	5.85 (1.17)	1.2 (0.9, 1.5)†
RAND Mental summary scale (0–100)			**Mean VAS**		
Nurse specialist patients	66.3 (24)	–8.6 (–14.5, –2.7)	Nurse specialist patients		73mm (+/–23)
Inpatients	53.0 (23)	–10.6 (–16.2, –5.0)†	Inpatients		85 mm (+/–19)
Day patients	51.3 (26)	–9.3 (–15.2, –3.3)*	Day patients		92mm (+/–10)*
					p <0.001

* Significant improvement between admission and week 52 p <0.05. † significant difference between day patient vs. inpatient adjusted for age, and differences at baseline p <0.01.

	HAQ: range of scores from 0 (no disability) to 3 (severe disability). A difference >0.22 has been found to be clinically relevant. MACTAR: interviewers assess which activities are most impaired and most important (maximum of 5) to the individual patient and follow the changes regarding these activities over time. RAND: high scores indicate better health. RAQoL: lower scores indicate better QoL. DAS: a composite index of disease activity. Patients' satisfaction was measured on a VAS ranges from 0 mm (dissatisfied) to 100 mm (satisfied). Over the total follow-up period all groups with respect to functional status, QoL, health utility and disease activity improved significantly over time (p <0.05). There were no sustained differences in clinical effectiveness between care provided by CNS and care provided by a multidisciplinary team either in an inpatient or a day care setting. Subgroup analysis using HAQ indicated the most favourable outcome for any type of care shifted from CNS and inpatient team in younger patients to day care in older patients.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 N/A 2 N/A 3 Yes, apart from patient satisfaction which was also reported at week 12 for the CNS group only 4 52 weeks. Loss of follow-up: 5 died and 21 lost due to either severe comorbidity, deteriorating physical condition, unwillingness or removal. 5 Yes, stratified by gender by an independent investigator 6 6 academic and non-academic hospitals in Leiden
Commentary	Despite the fact that the randomisation procedure was executed in blocks and by an independent investigator, the three groups were not completely comparable at baseline, so results may be affected by confounding factors. CNS patients were significantly younger, had better scores on the HAQ, RAND-36, and RAQoL questionnaires than day patients and inpatients. More CNS patients were employed than day patients or inpatients. CNS patients had significantly lower disease activity than day patients. This study was performed in patients whose condition made it acceptable for them to be randomised to all three types of care. Keeping this in mind, factors that may eventually play a role in the choice of treatment of patients with RA and functional limitations may be apart from age, the presence of complications and comorbidity, the availability of multidisciplinary facilities, patients' and doctors' preferences, and financial considerations.
Research implications	Qualitative analysis is required to examine in which ways and to what extent the different types of care and meeting patients' needs regarding individual functional limitations in various age groups are needed.

ID, origin, authors (year)	156, Netherlands, Vrijhoef, H.J.M., Diederiks, J.P.M., Spreeuwenberg, C. *et al. (*2001)
Aims	Assessment of effects on quality of care, in terms of patient outcomes, when tasks in the care for outpatients with stable type 2 diabetes are transferred from internist to nurse specialist and from outpatient clinic to general practice *Workforce:* Nurse specialist (NS), registered with the highest level of qualification, specialised in diabetes and with long-term work experience; primary care *Feature:* Substitution of internist for NP and transfer from outpatient clinic to general practice, for care of outpatients with stable type 2 diabetes *Outcome:* Glycated haemoglobin concentration (HbA_{1c}) fasting total cholesterol, HDL-cholesterol, triglycerides, body mass index (BMI), systolic (SBP) and diastolic blood pressure (DBP), health status, self-care behaviour, knowledge of diabetes, patient satisfaction and number of consultations with care providers
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Nonequivalent control group 2 Inclusion criteria: diagnosed with noninsulin dependent diabetes mellitus (WHO criteria), glycated haemoglobin (HBA_{1c}) <10.5% for the preceding 6 months at least. Measurement of HBA_{1c} has to take place at three different and successive moments in time. The most recent concentration has to be within the range of 1%-point of the mean for all three measurements. Exclusion criteria: presence of active complications (micro- and macro-angiopathy), presence of other diseases not related to diabetes mellitus for which care of a medical specialist is received, presence of psychosocial problems (assessed by physician). 3 121 patients with type 2 diabetes: 74 in the intervention group, 47 in the control group. 4 Dates not specified 5 Clinical notes; COOP/WONCA charts; visual analogue scale (VAS); self-care behaviour checklist (SCBC); Dutch diabetes-specific instrument; patient satisfaction questionnaire, over a 12 month period
Results Quantitative results	*Changes in mean HbA_{1c} for the intervention subgroup treated with OHA and/or insulin and the control group*

Group		**0 months**		**12 months**		**Within group**		**Between group**	
	n	***Mean, SD***		***Mean, SD***		***F-statistic***	**p-*value***	***F-statistic***	**p *-value***
Intervention subgroup	52	8.3	1.5	8.2	1.0	3.776	0.012*		
Control	46	8.2	1.1	8.5	1.4	2.744	0.044*	5.999	0.000*
With complete data:									
Intervention subgroup	31	8.6	1.4	8.3	1.0	3.396	0.018*		
Control	23	8.6	1.1	8.8	1.3	2.243	0.099*	5.386	0.001*

* Greenhouse-Geisser adjusted univariate approach

Outcome	**Intervention subgroup**	**Control**
Mean total cholesterol	Declined 0.5 mmol/l*	No difference
Mean HDL-cholesterol	Increased 0.1 mmol/l*	No difference
Triglycerides	No difference	No difference
BMI	No statistically significant change within or between groups	
SBP	Increased 3.6 mmHg*, no change in mean SBP between groups	Decreased 3.0 mmHg *
DBP	No changes between or within groups	
Mean satisfaction mark	7.8 ¯1.4 n = 29 No significant changes within or between groups	8.1 ¯1.0 n=21

* Statistically significant

Effects on self-care and knowledge for the intervention subgroup treated with OHA and/or insulin and the control group

		Measurement: Mean and SD		Within group		Between group	
Outcome (min–max)	***n***	***0 months***	***12 months***	***F-statistic***	**p-*value***	***F-statistic***	**p-*value***
Quality of life							
Physical fitness (5–1)							
• Intervention subgroup	31	3.3 ± 1.5	3.2 ± 1.4	0.390	0.679		
• Control	20	3.2 ± 1.3	3.2 ± 1.3	0.110	0.896	0.281	0.720
Feelings (5–1)							
• Intervention subgroup	31	2.3 ± 1.2	2.2 ± 1.0	0.498	0.610		
• Control	23	2.0 ± 1.2	2.3 ± 1.1	1.335	0.272	1.107	0.335
Daily activities (5–1)							
• Intervention subgroup	32	2.4 ± 1.1	2.7 ± 1.2	1.239	0.297		
• Control	23	2.0 ± 1.1	2.2 ± 1.2	0.702	0.501	0.935	0.396
Social activities (5–1)							
• Intervention subgroup	29	2.1 ± 1.3	2.0 ± 1.1	0.330	0.721		
• Control	23	1.9 ± 1.0	2.0 ± 1.2	0.169	0.787	0.391	0.678
Change in health (5–1)							
• Intervention subgroup	30	2.9 ± 0.7	2.9 ± 0.9	0.021	0.980		
• Control	23	2.9 ± 0.5	2.9 ± 1.0	0.117	0.890	0.096	0.908
Overall health (5–1)							
• Intervention subgroup	32	3.4 ± 0.8	3.6 ± 0.7	1.130	0.322		
• Control	23	3.6 ± 1.0	3.3 ± 0.9	0.870	0.391	1.951	0.147
VAS (0–10)							
• Intervention subgroup	30	4.3 ± 2.5	4.3 ± 2.8	0.009	0.991		
• Control	21	4.2 ± 2.6	4.2 ± 2.9	3.076	0.057	1.324	0.271
Self–care behaviour							
Diet application (1–5)							
• Intervention subgroup	19	3.2 ± 1.1	3.4 ± 1.0	1.332	0.274		
• Control	15	3.3 ± 0.7	3.3 ± 1.0	0.432	0.593	0.939	0.374
Self–regulation (1–5)							
• Intervention subgroup	30	3.1 ± 1.2	3.6 ± 1.1	4.617	0.014		
• Control	22	3.9 ± 1.0	4.0 ± 1.0	0.256	0.776	1.403	0.251
Activity of condition (1–5)							
• Intervention subgroup	19	3.0 ± 1.3	3.0 ± 1.1	0.010	0.990		
• Control	16	2.7 ± 1.2	3.0 ± 1.2	0.899	0.389	0.418	
Overall (1–5)							
• Intervention subgroup	18	3.2 ± 0.8	3.5 ± 0.7	1.526	0.236		
• Control	13	3.6 ± 0.6	3.5 ± 0.6	0.370	0.694	0.696	
Knowledge (0–12)							
• Intervention subgroup	32	7.7 ± 3.3	8.0 ± 3.0	0.451	0.639		
• Control	23	8.3 ± 2.7	8.7 ± 2.3	1.125	0.334	0.245	

	Outcome (min–max) **Group**	**n**	**0 months**	**12 months**	**Within group (Chi square, *p*-value)**	**Between groups (Chi square, *p*-value)**	
						0 months	***12 months***
	Consultations with NS						
	• Intervention subgroup	31	0.6 ¯1.1	2.4 ¯1.4	33.146, 0.000		
	• Control	22	0.8 ¯1.5	0.8 ¯1.1	0.333, 0.846	0.170, 0.680	16.919, 0.000
	Consultations with GP						
	• Intervention subgroup	31	0.5 ¯1.3	1.0 ¯1.4	1.962, 0.375		
	• Control	22	0.98 ¯1.2	1.2 ¯3.4	4.688, 0.096	1.286, 0.257	2.274, 0.132
	Consultations with internist						
	• Intervention subgroup	31	2.12 ¯1.6	0.9 ¯0.5	26.248, 0.000		
	• Control	22	1.9 ¯0.8	1.9 ¯0.9	0.047, 0.977	0.563, 0.453	18.141, 0.000
	Consultations with NS + GP + internist						
	• Intervention subgroup	31	3.2 ¯2.1	4.3 ¯2.2	4.750, 0.093		
	• Control	22	3.4 ¯2.2	3.9 ¯4.0	3.233, 0.199	0.051, 0.822	3.757, 0.053
	This study provides preliminary evidence proving consistent follow-up care using NSs who follow protocols and who are the primary interface with the patient, is an appropriate solution for managing the care of stable diabetic outpatients.						
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 None 2 None 3 Yes 4 Data from all questionnaires were available for 54 patients, with 59.6% and 48.9% available for the intervention subgroup and the control group respectively. 5 No 6 22/82 General practitioners (GP) referring patients to one university hospital in the Maastricht region						
Commentary	The optimal design to tackle the issue of causal inference is the randomised control trial (RCT). The requirements for GPs to participate in this study did not allow random allocation. GPs who participated in the substitution model may have a special interest in innovations with respect to diabetes care. General instruments were used to enable assessment of effects of the substitution model more generally. Together with a research period of 1 year, this might have resulted in not finding existing effects in outcomes.						
Research implications	Improved glycaemic control might result from lifestyle changes and/or therapy. In this study no correlations were found between change in HbA_{1c} and change in dose of medication or change in self-care behaviour. However, stable diabetic patients, already familiar with nurse specialists, are transposed into restricted ranges of both level of HbA_{1c} and self-care behaviour. Although substitution of care providers occurred, no evidence was provided about the model's cost-effectiveness. The quantity of consultations consumed does not reflect costs.						

ID, origin, authors (year)	106, Netherlands, Vrijhoef, H.J.M., Diederiks, J.P.M., Spreeuwenberg, C. *et al.* (2002)
Aims	To evaluate the effects of a shared care model, with the diabetes nurse as main care-provider for patients with type 2 diabetes in a primary care setting, on patient outcomes *Workforce:* Nurse specialist; primary care *Feature:* Substitution from physician to diabetic nurse as main care provider within a shared care model for patients with type 2 diabetes *Outcome:* Change in glycated haemoglobin level (%HbA_{1c}), systolic and diastolic blood pressure, total cholesterol, HDL cholesterol, triglyceride, patient satisfaction, quality of life, self-care behaviour and disease-specific knowledge, consultation with care providers
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Quasi-experimental pre-test/post-test using a control group (referred to as outpatients) were drawn from a study carried out at Maastricht University to assess the effects on patient outcomes when tasks of diabetes care are transferred from endocrinologist to diabetes nurse and from outpatient to primary care. 2 Patients with previously documented type 2 diabetes attending the general practice 3 175 patients with type 2 diabetes 4 September 1997 and April 1998 5 Laboratory results; blood pressure readings; industrial marketing management; COOP/WNCA charts and Visual Analogue Scale (VAS); Dutch diabetes specific instrument; self-care behaviour checklist; clinical data questionnaires
Results Quantitative results	*Effects in glycaemic control % HbA_{1c}*

Group of patients	**n**	**First measurement (Mean, SD)**		**Last measurement (Mean, SD)**		***p*-value within group**	**Last – first measurement (Mean, SD)**		***p*-value between groups**
Shared care	158	8.0	1.5	7.7	1.3	0.001	–0.3	1.0	–
Complete data	98	7.6	1.2	7.4	1.0	0.069	–0.2	0.1	
Missing data	60	8.3	1.6	8.0	1.4	0.004	–0.3	0.8	0.511
Changed therapy	23	8.7	1.6	8.2	1.5	0.081	–0.5	1.3	
Unchanged therapy	135	7.7	1.3	7.6	1.1	0.010	–0.2	0.8	0.851
Shared care (OHA/insulin)	38	8.8	1.4	8.4	1.3	0.008	–0.4	1.0	
Outpatient care	46	8.2	1.1	8.5	1.3	0.005	+0.3	0.8	0.001

Outcome in shared care group	**Improved by**	**To**	**n**	***p*-value**
Mean diastolic blood pressure	4.0 mmHg	80.6 mmHg	124	0.000
Total cholesterol	0.1 mmol/l	5.6 mmol/l	130	0.048
Triglyceride	0.2 mmol/l	1.8 mmol/l	128	0.005
Mean systolic blood pressure complete data	+5.1 mmHg		80	0.016

NB: Due to insufficient data about blood pressure and lipids of outpatients, no analysis between the shared care group and outpatient group could be performed.
HDL-cholesterol values not reported.

Outcome (min–max)	n	Measurement (Mean, SD) *Baseline*		*Second post measurement*		*p*-value	*F*-statistic
Satisfaction							
Satisfaction rate (0–10)	95	8.0	1.3	8.1	1.0	0.308* Greenhouse-Geisser adjusted univariate approach	1.175
Quality of life							
VAS (1–10)	100	5.4	2.5	5.7	2.5	0.249	1.401
Physical fitness (5–1)	98	2.6	1.3	2.7	1.3	0.598	0.516
Feelings (5–1)	101	2.2	1.2	2.2	1.2	0.502	0.673
Daily activities (5–1)	101	2.0	1.1	2.1	1.0	0.569	0.565
Social activities (5–1)	100	1.8	1.1	1.9	1.0	0.286	1.260
Change in health (5–1)	100	2.7	0.7	2.8	0.8	0.525	0.647
Overall health (5–1)	102	3.1	0.9	3.2	1.0	0.671	0.399
Knowledge and self–care behaviour							
Knowledge (0–12)	103	7.1	3.3	7.9	3.1	0.000	8.799
Diet application (1–5)	73	3.5	1.0	3.4	0.9	0.766	0.267
Self–regulation (1–5)	47	3.4	1.2	3.9	1.2	0.012	4.619
Activity of condition (1–5)	76	2.9	1.3	2.7	1.2	0.007	5.141
Bodily observation and conditioning (1–5)	81	3.3	1.3	3.8	1.2	0.000	8.092
Overall self–care behaviour (1–5)	21	3.4	0.7	3.4	0.6	0.851	0.162

Outcome	n	Measurement (Mean, SD) *Baseline*		*Second post measurement*		Chi-square	*p*-value
Consultations with care-providers							
Consultations with diabetes nurse	87	0.6	1.3	1.7	1.7	60.316	0.000
Consultations with General practitioner (GP)	90	1.3	1.3	1.2	1.1	7.977	0.019
Consultations with endocrinologist	93	0.2	0.6	0.2	0.6	1.914	0.384
Consultations with diabetes nurse + GP + endocrinologist	85	1.9	1.9	3.1	2.3	45.452	0.000

Evidence from this study seems to justify the continuation of a model of shared diabetes care with the diabetes nurse as main care provider for patients with type 2 diabetes in a primary care setting.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 and 2 Diabetes treatment and duration of diabetes served as co-variables. Treatment was specified as diet only, diet with OHA, diet with OHA and insulin, and diet with insulin. 3 Yes 4 Complete apart from loss of follow-up from questionnaires, 25 at baseline, 53 at 6 months, and 72 at 12 months 5 No 6 Five general practices in the Venlo region
Commentary	Not all patients participated in the study or provided complete data. Response rates and completion rate were low. The study findings may have been biased because patients with incomplete data suffered, on average, longer with diabetes and had worse mean glycaemic control. Missing responses were estimated by using the last observed response or mean of the group value. The study was limited by the study design; a RCT would have been a more appropriate method. Some patients appeared to consult the diabetes nurse at baseline before the introduction of the new shared-care model.
Research implications	Applicability of findings to all patients with type 2 diabetes being treated in primary care requires further study. Future research should provide evidence about the cost-effectiveness of the shared-care model and the relation between diabetic complications and quality of life scores should be undertaken.

ID, origin, authors (year)	1161, UK, Whittington, Z., Cantrill, J., Hassell, K. *et al.* (2001)
Aims	To describe community pharmacy management of minor conditions after referral from one general practice *Workforce:* Community pharmacist (CP); primary care *Feature:* Substitution *Outcome:* Prescribing outcomes, referral back to practice rates, and repeat consultation rates for same problem
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Feasibility study 2 Patients at one GP office calling for treatment of constipation, cough, diarrhoea, dyspepsia, earache, hay fever, head lice, headache, nasal symptoms, sore throat, high temperature, and vaginal thrush. Patients were given the option to see the pharmacist or a nurse/doctor. 3 576 patients were seen by the pharmacists. 4 14 days 5 26 weeks beginning in August 1999. Data on pharmacy referrals came for the community pharmacy professional advice form for 6 months. Re-consultation rate data were collected from the GP computer system.
Results Quantitative results	21 patients were referred back to the practice. 33 patients re-consulted for the same minor condition within 14 days. *Outcomes of 'care the chemist' referrals (n=576)* **Outcome** — **Frequency** — **%** Saw CP and received advice and formulary medicine — 511 — 88.7 Referred to CP but did not attend — 27 — 4.7 Saw CP and 'rapid referred' to GP — 21 — 3.6 Saw CP and received advice only — 9 — 1.6 Saw CP and bought OTC medicine — 8 — 1.4 *Re-consultations after pharmacy visit* **Type of re-consultation** — **Within 3 days** — **Within 7 days** — **Within 14 days** — **Total** CP to CP — 4 — 6 — 9 — 19 CP to GP — 5 — 4 — 5 — 14 Total — 9 — 10 — 14 — 33
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 None 3 Yes 4 Not beyond 14 days 5 None 6 One general practice and 8 local pharmacies in Merseyside, UK
Commentary	There could be patient/volunteer bias. People who did not want to wait to be seen by a doctor or nurse may have additional socioeconomic factors that made seeing a pharmacist more attractive.
Research implications	Needs to be done on a larger scale with randomisation. Are there more problems that pharmacists could provide care for? How was patient satisfaction? Do patients trust the pharmacist to handle these situations?

ID, origin, authors (year)	225, Australia, Biro, M.A. *et al.* (2000)
Aims	To assess if a team midwifery model of care for low and high-risk women is associated with a lower rate of obstetric interventions, greater satisfaction for women and reduced length of stay postpartum than for the standard model of maternity care *Workforce:* Tertiary-level care centre setting *Nursing workforce:* midwife *Feature:* (a) Substitution of team midwife care for the standard care model (b) Team-work among midwives *Intervention/comparison:* Intervention care model is characterised by continuity of midwifery care from early pregnancy to early postpartum, and comprises antenatal clinic visits for low-risk women with midwives, apart from three scheduled visits with obstetric staff. High-risk women had an individualised care plan in consultation with an obstetrician, but obstetric clinic visits were also conducted with a team midwife present, and a team midwife usually provided intrapartum care within the protocols of the delivery suite, with team midwives providing up to a shift a day of care for team care women. Standard care model comprised a team of midwives who provided care at any stage in the pregnancy in collaboration with physicians. Options for standard care included shared care between general practitioners in the community and hospital obstetric staff, shared care between midwives in a health centre and hospital obstetric staff, care by hospital obstetric staff only, and care by hospital midwives in collaboration with obstetric staff, similar to antenatal team care. Variable levels of continuity of care were provided in these different care plans, but all were cared for by a variety of doctors and midwives. *Outcomes:* *Maternal* Analgesia by type; pudendal block; anaesthesia by type; no analgesia/anaesthesia; monitoring of labour by type; augmentation of labour; induction of labour; mode of delivery by type; perineal status by degree of tear; days in hospital *Infant* Admission to Special Care Nursery (SCN); reasons for admissions to SCN >5 days by type; total no. pre-term infants; Birthweight <10th centile for gestational age; Apgar scores <7 at 5 mins; perinatal deaths; days in SCN
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Low-risk and high-risk antepartum cases were involved in the trial. No baseline clinical characteristics were noted, nor any clinical criteria for inclusion/exclusion. 3 1000 women randomised, 502 to the team midwife care and 498 to standard care. 4 Follow-up began at booking, through antenatal care and birthing and until departed from the postnatal unit for mothers. For babies – not stated but presumed from birth until discharge from Special Care Nursery (SCN). 5 Data on interventions and maternal and infant outcomes extracted from hospital records and the hospital's computerised birthing database.

Results Quantitative results	**Maternal outcomes** *Delivery outcomes: team midwife care vs. standard midwife care – odds ratio (95% CI) unless otherwise stated* Analgesia (excluding elective Caesarean sections): nitrous oxide 0.94 (0.70–1.26); pethidine 0.74 (0.55–0.98); epidural 0.65 (0.47–0.90) Pudendal block: 0.95 (0.44–2.02) Anaesthesia: spinal 1.05 (0.70–1.59); general 1.44 (0.67–3.13) No analgesia/anaesthesia: 1.07 (0.71–1.60) Monitoring of labour (excluding elective Caesarean sections): continuous electronic fetal monitoring 0.72 (0.54–0.97); auscultation only 1.43 (1.06–1.91); no monitoring 0.71 (0.20–2.35) Augmentation (excluding elective Caesarean sections): 0.66 (0.48–0.90) Induction (excluding elective Caesarean sections): 1.19 (0.87–1.62) Mode of delivery: spontaneous 1.14 (0.86–1.51); operative vaginal 0.72 (0.50–1.04); emergency Caesarean section 1.41 (0.93–2.15); elective Caesarean section 0.76 (0.46–1.24) Perineal status (excluding all Caesarean sections): episiotomy 0.64 (0.46–0.90); sutured tear 1.16 (0.84–1.60); unsutured tear 3.54 (1.91–6.62); intact 0.82 (0.56–1.20) Days in hospital (mean): –0.3 (–0.05 to –0.04) **Infant outcomes** *Neonatal outcomes: team midwife care vs. standard midwife care – odds ratio (95% CI) unless otherwise stated* Admission to Special Care Nursery (SCN) 0.97 (0.69–1.37) Reasons for admissions to SCN >5 days: pre-term 0.39 (0.18–1.84); intra-uterine growth retardation (IUGR) 1.8 vs. 0 (undefined OR); birth asphyxia 0.00 (0.00–37.30) Total no. preterm infants: 0.83 (0.51–1.35) Birthweight: <10th centile for gestational age 0.92 (0.64–1.33) Apgar scores: <7 at 5 mins 1.17 (0.48–2.82) Perinatal deaths (20 weeks + gestation): 5 vs. 4 Days in SCN (mean): 2.0 (–5.6–1.7) **Summary** Overall, there were fewer procedures – epidural and narcotic use, augmentation of labour, electronic fetal monitoring, episiotomies – and higher rate of unsutured tears in the team midwife group. No statistical difference in operative vaginal deliveries, or overall Caesarean delivery rates. Continuity of midwifery care was associated with a reduction of medical procedures in labour and a shorter length of stay without compromising maternal and perinatal safety.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case-mix adjustment reported 2 None stated 3 There were gaps in data collection. Data were available from hospital records for 449 (of 502) women and 461 babies in the team midwife care group, and 439 (of 498) women and 452 babies in the standard care group 4 Intention to treat analysis regardless of loss to follow-up or withdrawal *Team midwife care:* 30 had miscarriage or termination, 14 lost to follow-up, 9 inadvertently re-recruited. Data available on 439 women. 2 sets of twins delivered. *Standard care:* 36 had miscarriage or termination, 18 lost to follow-up, 5 inadvertently re-recruited. Data available on 449 women. 13 sets of twins delivered. 5 Random allocation by computer 6 One medical centre in Victoria

Commentary	Continuity of care was the focus of this study but this may have been confounded by status of professional caregiver, although it could be methodologically difficult to assess. Demographic characteristics are reported but because there was no reporting of baseline clinical characteristics or any adjustment for case mix or regulation of clinical risk on entering the study using exclusion criteria, confounding exists when interpreting the clinical outcomes. From this study it appears that continuity of midwifery care could reduce involvement with medical procedures in labour and result in shorter length of stay.
Research implications	The reduced length of stay may have benefits that are worth exploring through an economic analysis. Future trials should use a recognised case-mix tool to allow for adjustment, or choose cases of only one type to reduce confounding, e.g. low-risk births only; complications of only one type; acuity of similar grades. A standard maternity case-mix tool may need development; however, research into how midwives manage more complicated cases may be restricted by ethical requirements.

ID, origin, authors (year)	258, USA, Bissinger, R. *et al.* (1997)
Aims	To look for a difference in length of stay, days on ventilator, days on oxygen, mortality, morbidity, and cost, when infants weighing between 500 and 1,250g are cared for by neonatal nurse practitioners versus medical house staff *Workforce:* Secondary care setting *Nursing workforce:* Neonatal nurse practitioners *Feature:* Substitution of neonatal nurse practitioners for medical house staff *Intervention/comparison:* To compare the outcomes of neonates under the care of neonatal nurse practitioners with the outcomes of neonates under the care of medical house staff in a 36-bed neonatal intensive care unit (NICU). Both medical house staff and nurse practitioners were supervised by the same lead physician . *Outcomes (infant):* length of stay in NICU; days on oxygen; days on ventilation; morbidity – frequency of sensorineural hearing loss (BAER), retinopathy of prematurity (ROP), intraventricular haemorrhages (IVH); mortality; quality of care index (calculation and weights provided). Costs of care for both groups were also recorded but were not presented in this abstract.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 *Inclusion:* All critically ill neonates admitted to NICU within first 24 hours of life during the study period and whose birthweights were between 500 and 1250g *Exclusion:* Infants admitted to NICU after the first 24 hours of life; infants who died within the first 24 hours of life; infants with congenital cardiac, genetic or surgical conditions 3 Final sample: 70 infants, 35 in each group. 35 infants in total fitted the criteria and were cared for by neonatal nurse practitioners. 187 infants cared for by medical staff fitted the criteria and 35 of these infants were chosen at random. 4 18 months 5 Computer database of admissions between 1 January 11991 and 31 July 31 1992. Information obtained from the database was verified by researcher from medical records of infants under study.
Results Quantitative results	**Infant outcomes** *Neonatal nurse practitioner vs medical house staff – days or % (*p-*value)* Mean length of stay in NICU (range): 43 (2–183) vs. 57 (6–229), p = 0.073 Mean days on oxygen (range): 21 (0–62) vs. 25 (1–106), p = 0.232 Mean days on ventilation (range): 29 (0–97) vs. 40 (2–218), p = 0.097 *Morbidity score:* BAER p = 0.87 Pass 86 vs. 84 Fail 14 vs. 16 ROP p = 0.17 Present 19 vs. 0 Absent 81 vs. 100 IVH p = 0.30 Normal 68 vs. 56 Grade I 12 vs. 6 Grade II 3 vs. 19 Grade III 17 vs. 13 Grade IV 0 vs. 6 *Mortality:* 20 vs. 14 p = 0.53 *Quality of care index score:* 1.01 vs 1.02, where a value of 1.00 suggests average quality

	Summary There were no significant differences in the outcomes of care. The quality of care index was similar for both nurse practitioners and medical house staff, and neonatal nurse practitioners appear to be an acceptable alternative care provider to medical house staff.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 Results based on sample of 68 infants due to large differences in standard deviations between nurse practitioner and medical staff infants for two of the outcome variables. On examination of the records two infants cared for by medical house staff were removed from the analysis. 3 No gaps in data collection reported. 4 Records of infants studied from admission until death or discharge. 5 Random sampling of infants under medical staff care. 6 One regional referral neonatal intensive care unit in a university hospital
Commentary	Patient characteristics were reported and the groups were similar. Medical house staff rotate through the unit too frequently for them to follow up an admission for more than one month, similarly for the attending physician. This leads to a lack of continuity and fragmentation of care and nurse practitioners may be able to fulfil this role better than medical house staff.
Research implications	Future research should focus on the process and contextual variables that can influence the provider behaviour and ultimately the outcomes of care. While it is easy to compute the dollar value for 0.01 increase in quality, future cost-effectiveness studies should be accompanied by a sensitivity analysis that computes the cost-effectiveness ratio for specific outcomes when assigned different weights, because it is more difficult to evaluate the clinical value or impact of any change in the level of quality.

ID, origin, authors (year)	286, USA, Blanchette, H. (1995)
Aims	To compare the obstetric outcomes of patients in a primary care clinic under care of certified nurse-midwives (CNM) supervised by obstetricians with the obstetric outcomes of patients in a private practice under obstetricians' care *Workforce:* Primary care perinatal access clinic for indigent women and obstetric clinic Nursing: midwife *Feature:* Substitution of midwives for obstetricians *Intervention/comparison:* A comparison of supervised CNM care for medically indigent women and obstetrician care for private patients. 2 full-time and 5 part-time CNMs were supervised by a group of 4 full-time private obstetricians. *Outcomes:* *Maternal* Antepartum complications: late-entry pre-natal care; pre-term labour; small for dates; urinary tract infection; gestational diabetes; asthma; placenta praevia; labile hypertension; stillborn; twins Intrapartum and postpartum complications: postpartum haemorrhage; endometritis; retained placenta; amnionitis; abruptio placentae; pregnancy-induced hypertension; shoulder dystocia *Caesarean section* Caesarean section rate Indications for Caesarean section Previous Caesarean section *Fetal* Apgar score Birthweight
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Exclusion criteria not clearly stated. Patients excluded from the study were: diabetics (5), Rh-sensitised (1), cerebral aneurysm (1), chronic hypertension (2), congenital heart disease (1), twins with pre-term labour at 28 weeks (1), premature rupture of membranes with sepsis at 34 weeks (1). 3 Total sample size of 1107. 496 patients of the access clinic compared to 611 private patients. 4 Not stated. From onset of antenatal care until delivery. 5 Primary care access clinic and private practice records – data gathered over August 1991 to March 1994. Source of data not stated, but presumed to be maternity charts.

Results Quantitative results	**Maternal outcomes** *Maternal outcomes: number in midwife clinic (n=496) vs. number in obstetric care (n=611),* p-*value given where possible* Antepartum complications: Late-entry pre-natal care 67 vs. 22 (p <0.05); pre-term labour 9 vs. 8; small for dates 35 vs. 8; urinary tract infection 47 vs. 1 (p <0.05); gestational diabetes 4 vs. 5; asthma 8 vs. 8; placenta praevia 2 vs. 4; labile hypertension 4 vs. 3; stillborn 0 vs. 1; twins 1 vs. 13 Intrapartum and postpartum complications: Total: 47 complications (9.5%) vs. 13 complications (2.1%) Postpartum haemorrhage 11 vs. 5; endometritis 2 vs. 1; retained placenta 5 vs. 2; amnionitis 8 vs. 1; abruptio placentae 1 vs. 2; pregnancy-induced hypertension 4 vs. 2; shoulder dystocia 1 vs. 0 *Caesareans: midwife clinic vs. obstetric care % (p-value where significant)* Caesarean section rate (p <0.05) Primary 10.5 vs. 18.5 Repeat 2.6 vs. 7.9 Total 13.1 vs. 26.4 Indications for Caesarean section: dystocia 37 vs. 31.1; fetal distress 13.8 vs. 13.7; breech 13.8 vs. 11.8; previous Caesarean section 20.0 vs. 29.8 (p <0.05); other– 15.4 vs. 13.7 The clinic had a total Caesarean section rate of 13.1% vs. private practice Caesarean section rate of 26.4%; primary and repeat rates were also very low in the clinic. For complications leading to a Caesarean section, there was no significant difference except in previous Caesarean section. Of those with a previous Caesarean section, 18.2% private practice group attempted a vaginal birth, while 77.4% clinic group attempted a vaginal birth (p < 0.05). Successful vaginal birth rates after Caesarean section were similar in both groups. **Fetal outcomes** Apgar score: 1-min average 8.0 vs. 7.9; 7-min average 9.0 vs. 8.9; 1-min <7 (%) 8.0 vs. 9.66; 5-min <7 (%) 0.8 vs. 1.13 Birthweight (%): <5lb 2.4 vs. 3.07; 5–8 lb 71 vs. 66.88; 8–9 lb 20.3 vs. 21.65; >9 lb 6.3 vs. 8.4 **Summary** Midwives can successfully provide antenatal care and delivery with comparable fetal outcomes despite the substance abuse and late entry of a third of the clinic group into antenatal care. Apgar scores and birthweights were similar and there were no stillborns and maternal intrapartum and postpartum complications were low given the antepartum complications. Low-income, uninsured and under-insured women in supervised CNM care can have obstetric outcomes similar to women having prenatal care under obstetricians. Prenatal care with supervised CNMs can reduce the Caesarean section rate without compromising infant outcomes.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 None 3 Assumed – no gaps reported. 4 10 patients with medical problems transferred antepartum, and two patients transferred intrapartum with complications – these patients were not included in the study. 5 No sampling method – all records between chosen dates were analysed. 6 One private group obstetric practice with perinatal care access clinic in Berkeley.

Commentary	Demographics were well reported. Differences in the data were attributed to socioeconomic factors. The clinic group had a significantly later start to their prenatal care in the first and second trimester which may have skewed the results for antepartum complications and fetal outcomes. These differences were not adjusted for. Age differences were also statistically significant. The majority of the private group were white, whereas the majority of the primary care group were black or Hispanic.
Research implications	Certified nurse midwives have been attributed with skills in labour support that produce comparable outcomes between socio-demographically different groups, and there may be a place in perinatal care for CNMs to work with this group of women. To help describe the characteristics of the population of women in which this care can be put to best use, more prospective studies or studies of randomised controlled design would be useful in determining the contribution CNMs can make.

ID, origin, authors (year)	1176, USA, Butler, J. *et al.* (1993)
Aims	To test the hypothesis that supportive care during labour by a nurse–midwife would be associated with a lower incidence of Caesarean section in low-risk women. *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwives for physicians *Intervention/comparison:* Deliveries performed by midwives were compared to deliveries performed by physicians for the outcome of the method of delivery in low-risk women on a labour and delivery ward. Midwives provided additional one-on-one support for their patients in labour, depending on the patient's desire for support. The midwife stayed with the woman from early labour or until admission in active labour until after the birth and delivered one-on-one supportive care. Resident physicians in training could not spend as much time with labouring women as midwives could. Midwives consulted whenever problems arose. *Outcomes:* *Maternal outcomes* Outcomes of labour: Low or outlet forceps Midforceps Vacuum Labour epidural Blood transfusion Febrile morbidity Complications of labour: Abnormal labour Prolonged latent phase Active-phase arrest Slow slope active phase Occipitoposterior Deep transverse arrest Arrest of descent Prolonged labour Prolonged second stage of labour Fetal distress *Infant outcomes* 5-min Apgar score ≤7 Birth trauma Admitted to neonatal intensive care unit Small for gestational age Large for gestational age Birthweight

Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Inclusion: Women who were cared for by midwives or resident physicians who had at least 5 prenatal visits at the obstetric department; 37–42 week gestation; singleton; live-born; occiput presentation Exclusion: one or more medical or obstetric complications; congenital anomalies; elective repeat Caesarean section; uncertain care provider; more than one pregnancy in the study period Remaining indicators for Caesarean delivery were diagnoses of labour abnormalities or fetal distress. 3 Sample size of 4607: 1056 midwife patients, and 3551 physician patients 4 Records of patients were followed up from admission until delivery for mothers including admissions to neonatal intensive care unit for infants. 5 University of California, San Francisco Perinatal Database yielded data on labour, delivery and pregnancy outcomes as well as antenatal data. The records came from the period 1 January 1981 to 1 July 1988.
Results Quantitative results	**Maternal outcomes** *Maternal outcomes of labour: midwife vs. physician – adjusted odds ratio (95% CI)* Low or outlet forceps 0.06 (0.03–0.12); vacuum 0.04 (0.01–0.13); labour epidural 0.34 (0.28–0.42); blood transfusion 0.74 (0.27–1.98); febrile morbidity 0.77 (0.5-1.20) *Complications of labour: midwife vs. physician – adjusted odds ratio (95% CI)* Abnormal labour 0.70 (0.60-0.83): prolonged latent phase 1.04 (0.8–1.36); active-phase arrest 0.78 (0.60–1.02); slow slope active phase 0.74 (0.57–0.97); occipitoposterior 0.91 (0.64–1.28); deep transverse arrest 0.34 (0.15–0.77); arrest of descent 0.34 (0.26–0.45); prolonged labour 0.74 (0.56–0.99); prolonged second stage of labour 0.84 (0.68–1.03) Fetal distress 0.53 (0.32-0.77) **Infant outcomes** *Infant outcomes of labour: midwife vs. physician – adjusted odds ratio (95% CI)* 5-min Apgar score ≤7 1.11 (0.76–1.61); birth trauma 0.58 (0.33-1.04); admitted to neonatal intensive care unit 0.47 (0.24-0.95); small for gestational age 0.69 (0.49-0.96); large for gestational age 1.05 (0.81-1.35); birthweight 3512 g vs. 3429 g **Summary** The midwife sample had a lower incidence of Caesarean delivery than did physician patients – 9.75% vs. 12.3% (p = 0.02). After using multivariate analysis to estimate the risk of Caesarean section while controlling for the variables age, race, parity, year of delivery, and birth size, the association remained significant – adjusted OR 0.71 95% CI: 0.61–0.96). After adjustment for age, race, year of delivery, infant size and parity, the midwife sample experienced fewer abnormal labours, especially deep transverse arrest, arrested descent and prolonged labour. The midwife group also experienced fewer cases of fetal distress, a risk factor for Caesarean section. The physician sample were more likely to undergo an operative vaginal delivery with forceps and vacuum and were more likely to receive an epidural. There was a slight increase in the number of physician deliveries that were admitted to the neonatal intensive care unit, and the birthweight for midwife babies was higher.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 Multiple logistic regression was used to calculate the adjusted odds ratios treating year of delivery, socio-demographic characteristics and infant size as potential confounders. 3 No gaps in data collection reported. 4 Not specified. Details on transfers not provided. 5 Not applicable 6 One obstetric department at a university hospital, California
Commentary	Demographics were reported and the groups were dissimilar for race – a higher percentage of orientals, black and Hispanic women were found in the physician group and a higher number of white women were found in the midwife group. More nulliparous women were in the midwife group and more parous women were in the physician group. There was a statistically significant difference in the ages between the two groups although the authors believe this would not affect the outcome by increasing the risk of Caesarean section. The self-selection of women for one or other care provider is a major potential source of bias. The training of resident house officers may bias the results in favour of the more experienced midwives.
Research implications	Many research questions posed by this study have been answered in research since this study was published. The association of epidural anaesthesia with physician provider and increased risk of Caesarean section is probably not due to self-selection and may reflect a reduction in the requirement for such anaesthesia with midwife care – a future randomised controlled trial should include this association when setting the power. Fetal distress is associated with increased risk of Caesarean, and a further investigation into fetal distress and the factors associated with it should be done before studying the rate of occurrence of these factors in both midwife and physician patients.

ID, origin, authors (year)	1178, UK, Campbell, R. *et al.* (1999)
Aims	To compare the outcome of care given to women booked at a midwife-led unit with that for women booked at a consultant unit *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwife-led care for obstetrician-led care *Intervention/comparison:* The comparison of low-risk women who booked for delivery in the midwife-led maternity unit in one city and a similar group of women who booked for delivery in the consultant unit in another city. Most women who booked for delivery at the midwife unit booked with midwives, although they had been nominally placed under the care of a consultant obstetrician and others were booked for delivery under the care of their general practitioner. A domino scheme was also available. Care was received throughout the antenatal period from the care providers, the midwives consulting obstetricians as necessary. Outcomes: Maternal outcomes Labour and delivery: labour before 37 weeks; cephalic presentation; labour induced; labour augmented; anaesthesia/analgesia; delivery; length of labour State of perineum Blood loss over 500 ml Significant problems after delivery *Infant outcomes* Birthweight (g) Apgar score <7 Resuscitation required Congenital abnormalities Transferred to special care from delivery suite
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective cohort study 2 Inclusion criteria were those who satisfied the criteria for booking at the midwife unit Exclusion criteria: parity ≥5; multiparous and aged ≥38; primiparous and aged ≥35; height ≤ 5 feet; previous medical history – diabetes, cardiac disease, renal disease, deep vein thrombosis, pulmonary thrombosis; previous obstetric history – recent infertility, Caesarean section or hysterectomy, proven or suspected pelvic disproportion, rhesus antibodies, habitual postpartum haemorrhage, >2 previous abortions, previous stillbirth or neonatal death; previous gynaecological history – pelvic floor repair or myomectomy. 3 A sample size of 1499 in total with 794 under midwife care and 705 under obstetrician care comprising women who met the booking criteria. Transfers: 27.1% midwife group transferred to obstetrician group prior to labour; 12.7% midwife group transferred to obstetrician group during labour Approximately 60% women who booked to deliver with midwife unit actually delivered there. 4 From booking until delivery 5 Booking period – 1 November 1992 to 30 June 1993. Data for the study were abstracted by midwives from hospital records using a questionnaire form.

Results Quantitative results	**Maternal outcomes** *Labour and delivery: midwife group vs. obstetrician group (95% CI)* Labour before 37 weeks 4.6 vs. 4.3 (-1.76 – 2.44); cephalic presentation 96.4 vs. 95.6 (–1.5 – 2.4); labour induced 13.4 vs. 21.7 (–12.1 – –4.3); labour augmented 29.9 vs. 40.7 (–15.6 – –5.9) Anaesthesia/analgesia: None 14.8 vs 11.0 (0.5 – 7.3); nitrous oxide 61.5 vs. 69.4 (_12.7 – 3.1); pethidine 33.9 vs. 49.9 (–21.0 – –11.1); water bath 10.6 vs. 3.8 (4.2 – 9.4); TENS 4.7 vs. 6.1 (–3.1 – 0.9); epidural/spinal 13.7 vs. 19.8 (–9.9 – –2.3); general 3.8 vs. 2.8 (–0.8 – 2.8) Delivery: spontaneous 84.0 vs. 83.5 (–3.2 – 4.3); assisted 7.8 vs. 10.1 (–5.2 – 0.6); Caesarean 8.1 vs. 6.49 (–1 – 4.3) Length of labour (mean): first stage 272.3 vs. 304.7 (0.83 – 0.96); second stage 26.5 vs. 27.7 (0.85 – 1.08); frst and second stage 314.1 vs. 349 (0.84 – 0.96); third stage 7.1 vs. 6 (1.09 – 1.26) *Maternal outcomes: midwife group vs. obstetrician group (95% CI)* State of perineum: episiotomy– 16.8 vs. 24.6 (–12.0 – –3.7); tear 43.3 vs. 43.7 (–5.1 – –4.7) Blood loss over 500ml: 6.6 vs. 5.5 (–1.3 – 3.6) Significant problems after delivery: 10.7 vs 12.1 (–4.6 – 1.9) **Infant outcomes** *Infant outcomes: midwife group vs. obstetrician group (95% CI)* Birthweight (g) Under 1500: 0.5 vs. 0.3 1500–1999: 0.9 vs. 0.4 2000–2499: 2.4 vs. 3.8 Over 2500: 3.8 vs. 3.8 (–1.96 – 1.96) Apgar score <7 at 1 min: 11.9 vs. 16.3 (–8.2 – –1) at 5 min: 0.8 vs. 2 –2.4 – -0.02) Resuscitation required: 19.8 vs. 25.9 (–10.3 – –1.8) Congenital abnormalities: 2.4 vs. 1.9 (–0.9 – 2.04) Transferred to special care from delivery suite: 4.6 vs. 6.1 (–3.8 – 0.8) **Summary** The care received by women from the midwife unit was as safe and effective as the care received from the obstetrician unit. For the outcomes of labour and delivery, the midwife group was less likely to have an intervention of augmentation or induction. The midwife group was more likely to use no anaesthesia with fewer using nitrous oxide, pethidine, or an epidural or spinal anaesthetic. Women in the midwife group were less likely to have an assisted delivery and had shorter lengths of labour. There were fewer episiotomies in the midwife group, and the infant outcomes as measured by Apgar score and resuscitation were also better for the midwife group. There were similar rates of Caesarean section in each group.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 None reported 2 None reported 3 Concern that data was incomplete or fully reliable for stillbirths and neonatal deaths. 4 Follow-up was by intention-to-treat. 5 Not applicable 6 One hospital in Bournemouth and one hospital in Poole
Commentary	Characteristics were reported and the groups were similar. No socio-demographic variables were reported. The comparison was made between two different hospitals; therefore differences could exist that are a feature of the hospital care rather than a feature of the care provided by the health professionals. The study was reported with an annoying reference to the word 'booking'. The extent of midwife care is not fully reported – how much care was received antenatally from the care providers?
Research implications	This prospective cohort study found results that mirrored those of randomised controlled trials. This study design is prone to confounding but appears to produce valid results and may be organisationally more feasible than a randomised controlled trial. Larger trials are needed to detect mortality outcomes and attempts should be made to answer the question of the impact that midwife-led care can have on mortality. Of the other types of midwife-led care: they could be compared against each other to help determine which models were better, and which characteristics of the midwife-ed care are the most important in accounting for the lower rates of interventions and the higher rates of satisfaction.

ID, origin, authors (year)	1175, USA, Chambliss, L. *et al.* (1992)
Aims	To test the hypothesis that the low caesarean birth rate on the midwifery service was the result of patient selection bias *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwives for physicians *Intervention/comparison:* The management of low risk women in labour was compared between a group of women under midwifery management in a birth centre and a group of women under physician management on a traditional labour and delivery ward. Birth centre patients were managed exclusively by midwives unless consultation with physicians was sought, and women admitted to the birth centre were managed by an established protocol *Outcomes:* *Maternal outcomes* Labour and delivery outcomes Mode of delivery: normal spontaneous vaginal; Caesarean section; forceps; vacuum; total operative Length of labour: stage 1; stage 2; stage 3 Analgesia Oxytocin augmentation Episiotomy 3rd and 4th degree extensions *Infant outcomes* Apgar score <7: 1-minute; 5-minute Birthweight
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised blinded clinical trial, with blinding of participants and caregivers 2 Inclusion: age 16-–45 years; singleton vertex presentation; estimated gestational age of >36 weeks and <42 weeks; fetal weight >2500 g and <4000 g Inclusions also covered: previous Caesarean section by low transverse uterine incision; previous successful vaginal delivery after Caesarean section; class A1 non-insulin dependent diabetics with normal fasting glucose; spontaneous rupture of membranes with clear fluid, no meconium and uterine activity; women with no antenatal care and a haematocrit >30% Exclusion: oral temperatures ≥100 ºF; spontaneous rupture of membranes without labour; station –3 or higher; significant maternal or fetal complication (poorly controlled diabetes; hypertension; pre-eclampsia; fetal growth retardation 3 Sample size of 386 patients calculated as needed to detect a difference in Caesarean section rates. Final sample size 492; 487 patients were included in the study due to missing information for 5 patients: midwife group 229; physician group 234. 4 Admission in labour until postnatal discharge 5 Clinical outcomes were recorded from the patient's chart, delivery room logbook, computerised discharge summary.

Results Quantitative results	**Maternal outcomes** *Labour and delivery outcomes: midwives vs. physicians % (p-value) unless otherwise stated* Mode of delivery, n=253 vs. n=234: normal spontaneous vaginal– 229 vs. 234; Caesarean section 5 vs. 1 (NS); forceps 0 vs. 13; vacuum 0 vs. 5; total operative 5 vs. 19 (p = 0.006) Length of labour (hours and minutes): stage 1 10 hours vs. 8.4 hours (p = 0.02); stage 2 33 minutes vs. 45 minutes (p = 0.0005); stage 3 15 minutes vs. 20 minutes (p = 0.1) Analgesia: 10.5 vs. 23.8 (0.005) Oxytocin augmentation: 11.7 vs. 37.2 (0.0004) Episiotomy: 10.8 vs. 35.4 3rd and 4th degree extensions: 1.8 vs. 7.7 Fetal scalp electrode: 17.1 vs. 44.7 (p = 0.00005) **Infant outcomes** Apgar score <7 1-minute: 11 vs. 6 (NS) 5-minute: 1 vs. 0 (NS) Birthweight: 3400 vs. 3494 (p = 0.02). One infant in the physician group weighed 5100 g. **Summary** There was no statistically significant difference between the groups for Caesarean section rates. In comparing the operative vaginal deliveries, 18 of 235 women in the physician group had operative (instrumental) vaginal deliveries compared to 0 of 234 women in the midwife group. The midwife group had fewer episiotomies, and fewer third and fourth degree extensions. Operative vaginal deliveries were associated with these extended tears. Neonatal outcomes between the groups were similar, mean birthweight was higher in the physician group and there were no difference in Apgar scores. Physicians also used internal fetal scalp electrodes more often than midwives.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 No adjustment reported. 3 Gap in data collection reported for 5 patients whose information could not be found, and it was also recognised that clinicians were less likely to record the length of labour accurately. 4 Analysis was by intention to treat. 14 patients originally assigned to the midwives were transferred to the physicians – 9 had a spontaneous vaginal delivery, 5 had a Caesarean delivery. 5 Allocation to groups was unclear. Authors report that random assignment was performed using a sealed envelope but generation of allocation sequence not reported. 6 One university obstetric unit in California
Commentary	Patient characteristics were reported and the groups were similar. The generation of the allocation sequence is unclear and the method of assignment to groups was also unclear. Socio-demographic variables were not reported and the midwife group was likely to take in women who were at higher antenatal risk, although the population was classified as low risk. The authors concluded that there was no difference in the Caesarean section rates and that the previously low Caesarean section rate observed in the midwife birth centre was probably due to patient selection bias. This study should be repeated with women who are not just expected to result in spontaneous vaginal deliveries, but who are at low risk antenatally, and at low risk on admission in labour. The blinding of participants and caregivers may now be considered unethical, especially where signed consent was not obtained, as for this study.
Research implications	Proper randomised controlled trials with blinding of random allocation and assignment need to be conducted on a known low-risk population – both antenatally and at admission – in order for the question of how much influence a physician or midwife has over the outcome of delivery to be answered fully.A more standard definition of low risk according to universally agreed protocols is required to reduce the fluctuation in risk with the loose classification 'low risk'.

ID, origin, authors (year)	111, USA, Davidson, M. (2002)
Aims	To show: • the incidence of selected high-risk factors of the population cared for by certified nurse midwives (CNM) • the outcomes of the certified nurse–midwife (CNM) high-risk population • how they compare with a national sample *Workforce:* Secondary care setting Nursing workforce: midwives *Feature:* Substitution of midwives for other birth attendants *Intervention/comparison:* A comparison of risk factors and maternal and fetal outcomes of high-risk women cared for by a group of CNMs over ten years with those in a set of national data for one year *Outcomes:* The matched variables chosen as maternal and fetal outcomes were indicators of morbidity in both sets of CNM and national data *Maternal* Caesarean section; vacuum/forceps delivery; vaginal birth after Caesarean section; maternal fever *Infant* Meconium; Apgar score at 7 and 5 minutes of age *High risk factors matched with national data set:* Premature rupture of membranes; diabetes; pregnancy-induced hypertension; hydramnios; chronic hypertension; Rh sensitivity
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 803 cases were chosen because they had one or more high-risk factors, from a population of 5487. High-risk conditions included a mix of maternal medical and antenatal complications. Antenatal complications included premature rupture of membranes; pregnancy-induced hypertension; intrauterine growth retardation; hydramnios; birth defects and Rh sensitivity. Maternal medical complications included maternal drug use; diabetes; sexually transmitted diseases; hepatitis B; HIV; chronic hypertension; alcoholism; mitral valve prolapse; sickle-cell disease; acute psychiatric illness; lupus; epilepsy; Hodgkin's disease; thalassaemia. Other medical conditions were noted but occurred in a frequency of less than 0.1%. 3 803 high-risk women formed the CNM sample compared to a national sample of 3,891,494. A direct comparison with the 7% of the 1994 national dataset could not be made due to variances in the two samples. 4 Not stated. Report suggests period of follow-up began at booking, ending at delivery. 5 Maternity chart and delivery records provided the data for the CNM; midwifery clinical data were retrieved by a researcher from the delivery log for women during 1988 to 1998. For the national data set, natality statistics came from the Centers for Disease Control and Prevention dataset for 1994.

Results Quantitative results	**Maternal outcomes** Delivery outcomes: midwives vs. national dataset % Caesarean section 12.8 vs 20.7 Vacuum/forceps delivery– 3.6 vs 9.4 Vaginal birth after caesarean section (VBAC)– 73.5 vs 28.3 Maternal fever– 83.6 vs 68.9 **Infant outcomes** Infant outcomes: midwives vs. national dataset % Meconium: 12.2 vs. 5.8 Apgar score at 7 and 5 minutes of age: 1.7 vs. 2.5 Matched high-risk factors: midwives vs. all US births as % Premature rupture of membranes 18.6 vs. 2.9; diabetes 11.6 vs. 2.6; pregnancy-induced hypertension 7.2 vs. 0.03; hydramnios 4.9 vs. 0.01; chronic hypertension 2.1 vs. 0.01; Rh sensitivity 0.7 vs. 0.007 A statistical comparison was not made and only the percentages were presented. **Summary** CNM sample had better outcomes. 83% had a spontaneous vaginal birth, 4% higher than national average; 73.5% had a VBAC, 45.5% higher than national average. Instrument delivery rates and Caesarean section rates were also lower for CNM. The study indicates that CNMs can and do provide care to high-risk women with favourable outcomes with access to physician consultation and achieve higher spontaneous birth rates and fewer instrumental and Caesarean deliveries than the national average.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported 2 Matching for specified high-risk factors only. No other matching or statistical adjustment made 3 No gaps in the data reported 4 Not applicable 5 Not applicable 6 One non-profit hospital-based inner-city clinic, and a national sample
Commentary	Large data sets involved with unknown methods of data capture for the US national data set. Probability values and confidence intervals not presented due to methodological issues. The study involved a single site for the CNMs only and lacked a paired control group. Retrospective design limited analysis to certain variables only. Demographics were not reported, but the range and frequency of high-risk conditions that could be found in the sample of CNM births was shown. The majority of women in the CNM sample were African American (98%) and used Medicaid (72%).
Research implications	To increase the generalisability to other populations, replication of the study design to other sites and settings is advisable since this population was distinctly socioeconomically different. An attempt should be made to use the most complete data with appropriate statistical adjustments – this time for low-risk births, since there are professional issues with midwives managing high-risk births

ID, origin, authors (year)	489, USA, Davis, L., *et al.* (1994)
Aims	To assess how medical interventions in labour impact on Caesarean section rates *Workforce:* Secondary private care setting Nursing workforce: midwife *Feature:* Substitution of midwives for obstetricians *Intervention/comparison:* The comparison was the impact of medical interventions in labour on maternal and neonatal outcomes of obstetrician with certified nurse–midwife managed (CNM) low-risk patients. *Outcomes:* *Maternal* Delivery methods: Caesarean section rate by parity; Caesarean section for failure to progress by parity; Caesarean section for fetal distress by parity; vaginal operative delivery by type Interventions: oxytocin; narcotic analgesia; epidural *Infant* Apgar score: <7 at 5 minutes Arterial cord blood pH: <7.0
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Records of low-risk women only were reviewed. Fetal and maternal complications increase risk for Caesarean births, and such records were eliminated from both CNM and physician groups. Risk factors were: gestational age less than 36 weeks; multiple gestation; malpresentation; placenta praevia; placental abruption; pre-eclampsia; diabetes; intrauterine growth retardation; chronic hypertension; cord prolapse; elective Caesarean section. A history of previous Caesarean section did not exclude cases if going to a vaginal trial of labour. All indigent clinical service patients were excluded because they were managed using separate protocols. 3 Population of possible cases: 573 CNM patients and 12,077 physician patients. Final cases: 529 CNM patients and 8,266 physician patients. Total: 8795. 4 4 years follow-up 5 4 CNM and 35 obstetricians involved in the management of the cases over years 1987–1990. Data on cases for mother and infants came from a perinatal database.
Results Quantitative results	**Maternal outcomes** *Delivery methods: midwife vs. physician % (*p-*value)* Caesarean section rate by parity: total 8.5 vs. 12.8 (p <0.004); primipara 12.7 vs. 18.1 (p <0.02); multipara 1.9 vs. 6.6 (p <0.007) Caesarean section for failure to progress by parity: total 7.9 vs. 11.3 (p <0.02*)*; primipara/FTP 12.4 vs. 15.8 (NS); multipara/FTP 0.1 vs. 6.0 (p <0.003) Caesarean section for fetal distress (FD) by parity: total 0.05 vs. 1.6 (NS); primipara/FD 0.3 vs. 2.3 (p < 0.02); multipara/FD 0.9 vs. 0.7 (NS) Vaginal operative delivery by type: total 5.3 vs. 17.1 (p = 0.0001); forceps 4.2 vs. 16.5 (p = 0.001); vacuum 0.01 vs. 0.6 (NS) Caesarean section after unsuccessful vaginal trial of labour: 5.0 vs. 23.9 (p < 0.05) *Interventions: midwife vs. physician % (*p-*value)* Oxytocin 32 vs. 53.6 (p = 0.0001*)*; narcotic analgesia 21 vs. 25 (p < 0.05); epidural 11 vs. 53 (p = 0.0001) **Infant outcomes** *Infant outcomes: midwife vs. physician % (*p-*value)* Apgar score: <7 at 5 minutes 0.95 vs. 0.53 *(NS)* Arterial cord blood pH: <7.0 0.5 vs. 0.3 *(NS)*

	Summary There was a statistically significant difference between the CNM- and physician-managed patients for both primiparas and multiparas for Caesarean section rate. Indications for operative delivery were most often failure to progress – the number of women undergoing caesarean section for this indication was significantly lower in the CNM group. Vaginal trial of labour was significantly more successful in CNM. Use of interventions – oxytocin, augmentation and epidural – were significantly higher in physician-managed women. Use of instrumental deliveries – forceps and vacuum – were significantly higher in physician-managed women. **Provider – CNM or physician – was not found to be significant for predicting Caesarean section following a regression analysis, although associated factors were significantly higher in incidence in the physician group**
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported. 2 None reported. 3 No gaps in data collection reported. 4 Retrospective case review – no gaps to follow-up reported. Presumed data collected from booking visit until postpartum discharge. 5 Patients choose care provider. 6 One women's hospital in Chicago
Commentary	Patients choose care provider so self-selection may influence the outcomes. Also, indigent women were left out of the study because they were involved in another study; therefore the lower socioeconomic classes are not represented, limiting the generalisability of the results. Retrospective data were used which has its own limitations over time and with regard to accuracy. Basic demographics were reported and the groups were similar except for race.
Research implications	This result could have important implications for policy. The management of low-risk births by midwives may be just as or even more successful than for physicians, with fewer Caesarean sections and therefore fewer resources being used in the midwife group. This study should be repeated over a number of sites to help determine the validity of the results, preferably using a controlled trial or prospective cohort design. An economic analysis could also be conducted to estimate the actual difference in resource utilisation.

ID, origin, authors (year)	267, Canada, Harvey, S. *et al.* (1996)
Aims	To determine whether midwifery care is as effective as medical care for low-risk women with respect to clinical outcomes *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of nurse-midwife care for obstetrician or family practice physician care Intervention/comparison: The intervention comprised the provision of care throughout labour, delivery and the immediate post-partum period by a team of seven nurse–midwives. Women in standard medical care had a choice of family physician or obstetrician as normal. This study was conducted on a pilot nurse–midwifery service. *Outcomes:* *Maternal* Obstetric interventions: Caesarean section; episiotomy; epidural for labour; ultrasound; biophysical profile; dietary supplement; amniotomy; induction of labour; labour augmentation; intravenous in labour *Antenatal complications* Bleeding before 20 weeks; bleeding after 20 weeks; placenta praevia; abruptio placentae; gestational diabetes; mild pregnancy-induced hypertension; herpes; urinary tract infection; pyelonephritis; influenza; decreased fetal movement; postdates; pre-term labour; malpresentation; large-for-gestational age suspected or confirmed; small-for-gestational age suspected or confirmed; polyhydramnios *Postpartum complications* Postpartum haemorrhage; retained placenta; temperature >38 ºC; severe haemorrhoids Lacerations: 1st degree; 2nd degree; 3rd degree; 4th degree; labial; periurethral; vaginal
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial, unblinded random allocation of participants 2 Detailed exclusion criteria not given. Eligible participants described as requesting midwifery care, at low-risk for antenatal complications according to the Alberta perinatal risk scoring system, and who were able to give informed consent. Women were excluded if they were primigravidas under 17 or over 37, had undergone a previous Caesarean section, or were >20 weeks gestation at the time of consent. 3 Sample size of 218 in total, 109 randomised to midwife care, and 109 randomised to physician care 4 From point of booking until 6 weeks post-partum for women 5 Data were collected using the Nurse-Midwifery Clinical Data set (Div. Research of Am. Coll. Nurse–Midwives) and adapted for the Canadian setting. Data collected on neonatal morbidity, maternal morbidity, and intervention rates.

Results Quantitative results	**Maternal outcomes** *Obstetric interventions: Nurse–midwife care vs. physician care % (95% CI), p-value reported where possible* Caesarean section 4.0 vs. 15.1 (*p* = 0.01; 2.89 to 19.3); episiotomy 15.5 vs. 32.9 (*p* = 0.007; 4.85 to 30.1); epidural for labour 12.9 vs. 23.7 (–0.044 to 27.6); ultrasound 58.4 vs. 80.6 (*p* = 0.001; 9.7 to 34.8); biophysical profile 22.8 vs. 26.9 (–8.06 to 16.3); dietary supplement 22.8 vs. 62.4 (26.8 to 52.4); amniotomy 16.8 vs. 30.1 (1.44 to 25.1); induction of labour 8 vs. 15.6 (–1.84 to 16.1); labour augmentation 14 vs. 21.1 (–4.04 to 17.2); intravenous in labour 26.7 vs. 42.9 (1.97 to 28.4) *Antenatal complications: nurse–midwife care vs physician care %* Bleeding before 20 wks 3 vs. 4.3; bleeding after 20 weeks 1 vs. 1.1; placenta praevia 1 vs. 1.1; abruptio placentae 0 vs. 1.1; gestational diabetes 5.9 vs. 6.5; mild pregnancy-induced hypertension 4 vs. 4.3; herpes 0 vs. 3.2; urinary tract infection 6.9 vs. 2.2; pyelonephritis 0 vs. 3.2; influenza 3 vs. 1.1; decreased fetal movement 2 vs. 1.1; postdates 4 vs. 5.4; pre-term labour 1 vs. 4.3; malpresentation 2 vs. 4.3; large-for-gestational age suspected or confirmed 2 vs. 1.1; small-for-gestational age suspected or confirmed 4 vs. 2.2; polyhydramnios 0 vs. 1.1 *Postpartum complications: nurse–midwife care vs. physician care %* Postpartum haemorrhage 5.9 vs. 3.2; retained placenta 2.9 vs. 2.2; temperature >38 ºC 1 vs. 2.2; severe haemorrhoids 0 vs. 2.2 lacerations, 1st degree 19.5 vs. 24.5; 2nd degree 28 vs. 37.7; 3rd degree 0 vs. 0; 4th degree 0 vs. 0; labial 3.7 vs. 3.8; periurethral 9.2 vs. 0; vaginal 9.2 vs. 0 *Other maternal outcomes:* Spontaneous vaginal delivery rate (%) 88.2 in nurse–midwife care vs. 76.3 in physician care Instrumental vaginal delivery rate (%) 5.9 in nurse–midwife care vs. 7.6 in physician care Length of stay postpartum (hours) 21.77 in nurse–midwife care vs. 51.68 in physician care (*p* = 0.0001; 95% CI, 22.4 to 38.2) **Infant outcomes** Apgar scores of <7 at 1 minute (n) 14 in nurse–midwife care vs. 27 in physician care (*p* = 0.013; 95% CI, 3.75 to 26.6) Apgar scores of <7 at 5 minute (n) 4 in nurse–midwife care vs. 4 in physician care Transfer to special or intensive neonatal care unit (n) 8 in nurse–midwife care vs. 18 in physician care (*p* = 0.02; 95% CI, 1.8 to 21) Average birthweight (g) 3502 for infants in nurse–midwife care vs. 3492 for infants in physician care (*p* = 0.886; 95% CI, 150 to 130) **Summary** Women in the nurse–midwife group experienced significantly fewer interventions in the intrapartal period – fewer intravenous infusions, amniotomies, and episiotomies than for physicians. Women in the nurse–midwife group had fewer induced or augmented deliveries (not significant). The non-interventionist style of care of nurse–midwives is due to more selective use of technology and interventions. Nurse–midwives can provide safe and effective care for women, showing fewer applications of technologic assessment, fewer interventions, shorter hospital stays, fewer neonatal intensive care unit admissions and less maternal morbidity than for equally low-risk women under care of physicians.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 None stated. 3 Data collected using the tool described above – the Clinical Data Set was chosen because it addressed three components of quality assurance: structure, process and outcomes. 4 Participants were followed up by intention to treat. There were 24 attritions, with 8 from the nurse–midwife group and 17 from the physician group. A complete intention to treat analysis was not possible due to the trial protocol; 4 from each group experienced spontaneous abortions after randomisation, 1 from each group failed to meet the inclusion criteria on further testing, 4 women rejected physician care and failed to meet the inclusion criteria, 1 in the nurse–midwife group and 3 in the physician group moved out of the research area, 4 selected physicians refused to co-operate, 2 withdrew from the nurse–midwife group. 5 Random allocation using consecutively numbered opaque envelopes 6 One midwife unit in a Calgary hospital for the intervention, and other city hospitals for the standard

Commentary	The exclusion criteria are not clear – no example of the Alberta perinatal risk-scoring system is provided, but the baseline risk is presumed to be low since only low-risk women were invited into the study. The authors say cultural differences could exist between the secondary care hospitals involved. Low risk was assessed in a self-report questionnaire; therefore women found to be high risk at the first antenatal clinic were taken out of the study. Women choosing home births were also excluded, and physicians who refused to co-operate with the study protocol also took their patients with them. The problems of volunteering for studies include the desire for active participation. The location in a tertiary medical centre may increase the rate of use of interventions over that of a similar model practised in a less well-equipped centre. The demography is well reported and the groups were similar. Results show safe and effective maternity care can be delivered by nurse–midwives for low-risk mothers.
Research implications	Further studies of this design and cost-effective analyses of this maternity care model would help the Canadian health system to recognise the value of midwifery. Multi-site study designs with the adoption of the pilot intervention across different sites would help to establish the efficacy of the intervention in different regions, although cultural differences may influence practice and outcomes and these need to be approached sensitively.

ID, origin, authors (year)	292, UK, Hundley, V.A. *et al.* (1994) – linked to study 908
Aims	To look for differences between midwife-managed and consultant-led care in the intrapartum care and delivery of low-risk women in terms of maternal and perinatal morbidity *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwife-managed care for consultant-led care Intervention/comparison: The intervention was the intrapartum care and delivery care received in a midwife-managed delivery unit and the control was standard care on a consultant-led labour ward. Clinical outcomes of low-risk women assigned to these groups were examined. *Outcomes:* *Maternal* Outcomes of labour: Mode of delivery by type; state of perineum by tear severity; third-degree tear; mean estimated blood loss; placental delivery by intervention; mean length of postnatal stay Events during labour: Onset by type; augmented labour; mean gestation; mean length of labour by stage; delay in labour by stage; use of monitoring by type; use of fetal scalp electrode; analgesia by type; mobility; maternal/fetal complications by type *Infant* Fetal outcomes: Mortality by type; loss of pregnancy; lost to follow-up; mean birthweight; median Apgar score; mean pH of cord; resuscitation by type; no babies admitted to neonatal unit; mean length of stay in neonatal unit *Other outcomes recorded but not reported in this abstract* Maternal/fetal complications necessitating antepartum transfer off midwife unit Maternal/fetal complications necessitating intrapartum transfer off midwife unit
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Low-risk mothers only took part. Exclusion criteria: pre-existing maternal disease, infertility, previous complicated obstetric history (Caesarean section, difficult vaginal delivery, or poor obstetric outcome), height <150 cm, maternal age >35 years, multiple pregnancy 3 Allocation was 2:1 in favour of midwives' unit due to the expected rate of transfer to the consultant-led unit. 2844 expectant mothers were randomised to the midwife unit (n=1900) or the labour ward (n=944). 4 Involvement in study began at trial booking period, until delivery for the clinical aspect. (Questionnaire follow-up began at prepartum admission ending at postpartum discharge – questionnaire survey presented in study 908.) 5 Six sources used to collect information: (i) Staff questionnaire, completed by midwife after delivery; (ii) client questionnaire, completed by woman after discharge; (iii) interviews of random sample of 400 participants; (iv) case-note review; (v) Scottish Morbidity Register forms; (vi) Aberdeen maternity and neonatal databank. Trial booking period October 1991 to December 1992. Note: not all outcomes from these data sources are reported in this study or abstract. The expectations, experiences, satisfaction of parturient women, plus the role, experiences, and satisfaction of midwifery staff and costs of care are reported elsewhere in study 908 (Hundley *et al.*, 1997).
Results Quantitative results	**Maternal outcomes** *Outcomes of labour: midwives unit vs. labour ward % (*p*-value) unless otherwise stated* Mode of delivery: Spontaneous vaginal delivery 78.2 vs. 75.3 (p = 0.5); vaginal breech 1.3 vs. 1.3 (p = 0.5); forceps/ventouse 12.2 vs. 13.3 (p = 0.5); emergency section 6.9 vs. 8.0 (p = 0.5); elective section 1.5 vs. 2.1 (p = 0.5) State of perineum: intact 23.7 vs. 20.9 (p = 0.8); episiotomy 25.2 vs. 29.1 (p = 0.04); tear 51.1 vs. 50.1 (p = 0.8); third-degree tear 0.8 vs. 0.3 (p = 0.1) Mean estimated blood loss: 156 vs. 163 (p = 0.1) Placental delivery: controlled cord traction 94.7 vs. 95.6 (p = 0.6); maternal effort 3.1 vs. 2.4 (p = 0.6); manual removal, no anaesthetic 0.8 vs. 0.4 (p = 0.6); manual removal with anaesthetic/epidural 1.5 vs. 1.6 (p = 0.6) Mean length of postnatal stay (days): 2.0 vs 2.0 (p = 0.2)

	*Events during labour: midwives unit vs. labour ward % (*p-*value) unless otherwise stated* Onset (p = 0.4); spontaneous 78.6 vs. 80.1; induced 21.4 vs. 19.9 Augmented labour 15.3 vs. 14.9 (p = 0.9) Mean gestation (weeks) 39.7 vs. 39.8 (p = 0.9) Mean length of labour by stage (hours): first stage 7.0 vs. 6.8 (p = 0.3); second stage 0.9 vs. 0.9 (p = 0.7) Delay in labour by stage: first stage 3.1 vs. 2.2 (p = 0.2); second stage 5.2 vs. 5.1 (p = 1.0) Use of monitoring by type: Pinard 30.2 vs. 15.1 (p = 0.001); Doppler 54.8 vs. 10.3 (p = 0.001); cardiotocograph 57.3 vs. 92.8 (p = 0.001) Use of fetal scalp electrode 26.1 vs. 31.9 (p = 0.001) Analgesia by type: none 1.9 vs. 1.8 (p = 0.9); natural methods 53.8 vs. 45.0 (p = 0.001); Entonox 84.1 vs. 83.3 (p = 0.6); TENS 34.5 vs. 27.4 (p = 0.001); pethidine/diamorphine 63.5 vs. 63.1 (p = 0.9); epidural/spinal 14.7 vs. 17.7 (p = 0.05) Mobility (p = 0.001); able to move most of the time 63.5 vs. 51.6; unable to move 36.5 vs. 48.4 Complications by type: fetal distress 18.5 vs. 22.4 (p = 0.02); meconium-stained liquor 13.8 vs, 14.1 (p = 0.9); pre-eclampsia 2.8 vs. 1.9 *(p = 0.2)*; pre-term delivery (<37 weeks) 2.6 vs. 3.0 (p = 0.6); shoulder dystocia 1.4 vs. 0.9 (p = 0.3); undiagnosed malpresentation 0.7 vs. 0.4 (p = 0.5); other 2.2 vs. 3.3 (p = 0.1) **Infant outcomes** *Fetal outcomes: midwives unit vs. labour ward % (*p-*value) unless otherwise stated* Mortality: live born 99.2 vs. 99.3 (p = 0.5); stillborn 0.3 vs. 0.4 (p = 0.5); neonatal death 0.5 vs. 0.2 (p = 0.5) Loss of pregnancy 2.4 vs. 1.8 (p = 0.4) Lost to follow-up 1.8 vs. 1.0 (p = 0.1) Mean weight of infant (g) 3427 vs. 3420 (p = 0.8) Median Apgar score: at 1 minute 9 vs. 9 (Mann-Whitney U test: 0.6); at 5 minutes 9 vs. 9 (Mann-Whitney U test: 0.5) Mean pH of cord 7.29 vs. 7.29 (p = 1.0) Resuscitation: none or mucus extraction only 79.4 vs. 82.6 (p = 0.05); Naloxone ± oxygen or IPPV 14.9 vs. 12.4 (p = 0.1); oxygen or IPPV only 5 .7 vs. 5.0 (p = 0.1) No babies admitted to neonatal unit: total 7.9 vs. 7.4 (p = 0.8); for up to 48 hours 1.3 vs. 1.4 (p = 0.8); for more than 48 hours 6.6 vs. 6.0 (p = 0.8) Mean length of stay (days) 3.5 vs. 3.3 (p = 0.8) **Summary** Women allocated to midwife unit were significantly less likely to have continuous electronic monitoring and more likely to have intermittent monitoring by Pinard or hand-held devices. Rates of interventions: operative vaginal delivery, induction and augmentation were no different statistically although the midwife group had fewer episiotomies. Significantly more women allocated to the midwife unit reported using natural methods of pain relief, while women on the labour ward were more likely to use epidural pain relief (not significant). Women in the midwife unit were significantly more likely to be able to move around for most of the time during labour. Significantly fewer women had an episiotomy on the midwife unit than for the labour ward. Women on the midwife unit had a lower rate of any type of intervention overall.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case-mix adjustment reported. 2 None stated. 3 No gaps in data reported. 4 Study groups were analysed by intention to treat. Midwife unit: 50% transferred to consultant-led ward; 4% withdrawn/lost to follow-up _ 46% received attention in midwife unit, n=870. Labour ward: 6% transferred off consultant-led ward; 3% withdrawn/lost to follow up _ 91% received attention on consultant-led ward, n=862. 5 Random allocation to groups using opaque consecutively numbered envelopes 6 A maternity hospital midwife-managed delivery unit in Aberdeen, Scotland

Commentary	No obvious limitations to this study were found. Demographics were reported and the groups were similar. Rigid exclusion criteria were given. Study confirms that midwife-managed care is as safe as consultant-led care, and authors suggest this option may be better for low-risk mothers. Authors also suggest continuity of care, satisfaction of mothers, satisfaction of staff and costs of care should be considered in planning services.
Research implications	More studies of this design are required to determine if these results can be repeated in different settings, e.g. community hospitals, where midwives conduct deliveries and for which direct access to the obstetric unit is not so readily available, as access to immediate on-site obstetric expertise could be a critical factor in the success of this intervention. Multi-site studies across a region are also useful in determining the acceptability of this intervention in different settings.

ID, origin, authors (year)	908, UK, Hundley, V. *et al.* (1997) – linked to study 292
Aims	1 To look for differences in satisfaction with care among midwife-managed patients compared to consultant-managed patients 2 To compare factors relating to continuity, choice and control between the two groups *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwife-led care for consultant-led care *Intervention/comparison:* A study comparing the satisfaction with intrapartum care and delivery of low-risk women assigned to either a midwife-managed delivery unit or standard care on a consultant-led labour ward. *Outcomes:* *Maternal:* satisfaction; continuity of carer; women's views of support during labour and delivery; choice; control Full results abstracted only where differences were significant.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Pragmatic randomised controlled trial, assignment was 2:1 in favour of the midwives' unit. 2 Selection criteria established those of low risk among women booking for delivery in general practitioner units. Criteria for establishing low-risk women: Exclusion criteria – pre-existing maternal disease, infertility, previous complicated obstetric history (Caesarean section, difficult vaginal delivery, or poor obstetric outcome), height <150 cm, maternal age >35 years, multiple pregnancy 3 2844 were randomised to the midwife unit (n=1900) or the labour ward (n=944) Midwife unit: 50% transferred to consultant-led ward; 4% withdrawn/lost to follow-up; 46% received attention in midwife unit, n=870 Labour ward: 6% transferred off consultant-led ward; 3% withdrawn/lost to follow up; 91% received attention on consultant-led ward, n=862 4 Follow-up was by questionnaire relating to hospital stay, from prepartum admission until postpartum discharge. 5 Women were recruited over a 14-month period voluntarily and by informed consent in 1991 to 1992. Client and staff data were collected by questionnaire. Client questionnaire relates to pregnancy, antenatal care, labour, delivery and also records demography. Questionnaires were used to collect data on satisfaction. Women who had suffered a perinatal death were not asked to complete questionnaires. Midwife staff also completed a questionnaire providing information on their staff details, their role, experience and satisfaction.
Results Quantitative results	**Maternal outcomes** *Satisfaction: midwives vs. obstetricians % (*p-*value) unless otherwise stated* Thinking back about what happened to you and what the staff did, how do you feel your labour and delivery were managed by the staff? (*p* = 0.02) • as you liked it in every way 78.1 vs. 73.4 • as you liked it in some ways but not in others 20.7 vs. 25.1 • not as you liked it at all 1.3 vs. 1.4 Satisfaction with overall experience (n) 8.0 vs. 8.0 (*p* = 0.1)

	*Continuity of carer: midwives vs. obstetricians % (*p-*value) unless otherwise stated* Approximately how many midwives looked after you while you were in the delivery unit? (n) 2.0 vs. 2.0 (p = 0.03) Staff that women reported seeing during either labour or delivery: • midwife– 97.1 vs 96.9 (p = 1.0) • hospital doctor 37.5 vs. 45.2 (p <0.001) • student midwife 62.0 vs. 48.3 (p <0.001) • student nurse 12.3 vs. 12.8 (p <0.8) • medical student 10.2 vs. 19.0 (p <0.001) • paediatrician 22.9 vs. 26.0 (p = 0.01) • anaesthetist 17.1 vs. 21.9 (p = 0.004) • other 1.0 vs. 3.2 (p <0.001) • don't know 3.9 vs. 5.1 (p = 0.2) *Women's views of support during labour and delivery: midwives vs. obstetricians % (*p-*value) unless otherwise stated* Chosen companion was present during labour/delivery? (NS) Present for how long? (NS) Did you feel there were lots of different people coming in and out while you were in labour ? (p = 0.003) • a lot 3.0 vs. 4.3 • quite a few– 16.3 vs. 21.1 • hardly any– 80.7 vs. 74.7 Were you and your companions left alone by the staff at any stage when it worried you to be alone? (p = 0.003) • no: 90.4 vs. 86.3 *Choice* Were you given any choice as to the way your baby's heartbeat was monitored? (p = 0.002) • yes 6.2 vs. 9.9 During labour, did you feel you wanted to move around or change position ? (p = 0.004) • yes 57.4 vs. 50.2 • no, not really 41.1 vs. 48.2 • don't know 1.2 vs. 1.6 Where women wanted to move, were they able to? (p = 0.007) • able to move most of the time 70.7 vs. 62.8 • unable to move 29.1 vs. 36.8 Did the hospital staff encourage you to move around and change position ? (p <0.001) • yes 75.3 vs. 66.9 Where the woman had a spontaneous vaginal delivery (SVD), would they have liked to have tried another position for delivery? (NS) Did you have any particular preferences about what happened in the third stage ? (p = 0.6) If you had a preference for the third stage did you get what you wanted? (p = 0.11)

	Control Labour management decisions *(NS)* Did you have any say in whether your waters were broken ? (NS) How was the decision made about the type of pain relief to us ? (p <0.001) • I made my own decision with the staff's approval 54.5 vs. 49.0 • I made my own decision with the staff's advice 0.9 vs. 1.2 • I was happy to follow the staff's advice 36.8 vs. 35.9 • The staff were insistent I take their advice and I couldn't refuse 0.6 vs. 1.4 • It all just happened and there was no decision made as such 0.8 vs. 0.8 **Summary** Most women expressed satisfaction with labour and delivery in the midwife unit group but it was not significant with no difference between groups in satisfaction with the overall experience.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 Bonferroni correction to reduce p-value from 5% to 0.1%. No other adjustment reported. 3 No gaps in data collection reported. 4 Participant follow-up by questionnaire: Midwife unit: 4% excluded due to loss to follow-up; 5% excluded for other reason; 91% sent questionnaire with a 97% response rate (1674) Labour ward: 3% excluded due to loss to follow-up; 8% excluded for other reason; 89% sent questionnaire with a 93% response rate (789) 5 Consecutively numbered opaque envelopes 6 A maternity hospital in Aberdeen, Scotland
Commentary	Assessing satisfaction is very difficult due to the nature of the outcome and the experience that is childbirth. The authors report that scales of measurement may lack specificity. Also the birth of a healthy normal child could offset reports of dissatisfaction. The aim was to measure satisfaction, and although a difference was found between the two groups it did not reach significance and the authors say it is possible that the scale used was too crude. Demographics were reported and both groups were similar. Linked to study by Hundley *et al.* (1994).
Research implications	More research is needed to show what factors are important to women if they are to have a positive childbirth experience, and how these factors are influenced, e.g. differences in geography, use of a team care approach etc. Certain issues surrounding measurement of satisfaction with childbirth need attention, especially the optimum time for measurement in the postnatal period. Measurement of satisfaction requires definitive research to help develop a tool that yields valid research material.

ID, origin, authors (year)	242, USA, Karlowicz, G. and McMurray, J. (2000)
Aims	To compare outcomes and charges of care delivery to extremely low-birthweight infants by neonatal nurse practitioners and paediatric residents *Workforce:* Secondary care setting Nursing workforce: neonatal nurse practitioners *Feature:* Substitution of neonatal nurse practitioners for paediatric residents *Intervention/comparison:* A study comparing the health outcomes of extremely low-birthweight infants (ELBW <1000 g) in a neonatal intensive care unit (NICU) cared for by neonatal nurse practitioners (NNPs) against those cared for by paediatric residents. NNPs and residents functioned independently of each other with no crossover, under the supervision of the same certified neonatologists. Admissions were assigned to either of the teams on an alternating basis by the charge nurse unless the acuity of one team census was higher than the other. The resident team comprised an attending neonatologist and 4 paediatric residents completing a 4-month rotation. The NNP team comprised 8.5 full-time equivalents drawn from 11 nurses who had 4 years' NICU experience plus nurse practitioner training and who were required to staff the unit around the clock. *Outcomes*: *Infant* Median length of stay; survived to discharge; severe IVH or PVL; threshold ROP; chronic lung disease (IVH – Intraventricular haemorrhage; PVL – periventricular leukomalacia; ROP – retinopathy of prematurity) Costs were also recorded but not presented in this abstract.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 All infants with birthweights <1000 g admitted to the NICU during the 2-year period between 1 September 1994 and 31 August 1996. Infants who died earlier than 12 hour of age, admitted after 1 week of age or with major malformations, chromosomal abnormalities or congenital infections were excluded. 3 230 infants were admitted to the unit and 29 were excluded due to major congenital malformation (n=8), major chromosomal abnormality (n=6), transfer (n=10), death before 12 hours age (n=5). Final study group population included 201 infants – NNPs cared for 94 infants and residents cared for 107 infants. 4 Follow-up time began at admission until survival at discharge. 5 Period of data collection 1 September 1 1994 to 31 August 1996. Data were taken from a neonatal database compiled by a research nurses abstracting clinical information from medical records.
Results Quantitative results	**Infant outcomes** *Infant outcomes – NNPs vs. residents% (*p*-value) unless otherwise stated* Survived to discharge: 76 vs. 77 (p = 0.87) Severe IVH or PVL: 27 vs. 18 (p = 0.17) Threshold ROP: 17 vs. 13 (p = 0.65) Chronic lung disease: 30 vs. 30 (p = 1.0) Median length of stay (days): 87 vs. 88 (p = 0.54); range (39–230) vs. (41–365) **Summary** No significant differences were found in major clinical outcomes for infants <1000 g , regardless of assignment to NNPs or paediatric residents. Neonatal nurrse practitioners and paediatric residents provided comparable care to extremely low-birthweight infants.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported. 2 None reported. 3 No gaps in data collection reported. 4 Deaths and transfers were excluded from analysis. 5 Alternate assignment 6 One NICU in an East Virginia teaching hospital
Commentary	Basic demographics and clinical characteristics were recorded and the groups were similar except for ethnicity. There is a degree of generalisability afforded due to the homogeneous population. The study is limited because it is not a randomised controlled trial and the supervision for the teams could not be quantified. Comparable outcomes mean more units could be encouraged to commit to staffing of neonatal intensive care units with neonatal nurse practitioners.
Research implications	Randomised controlled trials are needed to rule out selection bias in future studies, but only when sufficient studies of this type and design – prospective cohort studies or controlled trials – have been conducted and whose results suggest that outcomes between different health professionals are comparable and that there is no risk to the infants. The level of supervision needs to be quantified in future studies, and costs need to be reliably calculated.

ID, origin, authors (year)	703, USA, MacDorman, F. and Singh, G. (1997)
Aims	1 To look for significant differences in birth outcomes and infant survival between deliveries under certified nurse midwife care and physician care; 2 To show whether the differences exist after adjustment for medical risk factors and sociodemographic variables *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwives for physicians *Intervention/comparison:* A study comparing birth outcomes and infant mortality rates between births delivered by certified nurse midwives (CNMs) with all births of a national data set over one year *Outcomes:* *Infant* Comparison of CNM care in relation to physician care: infant mortality rate; neonatal mortality rate; post-neonatal mortality rate; low birthweight; mean birthweight Characteristics of physician- and CNM-delivered births also reported for all deliveries and for singleton, vaginal deliveries 35–43 weeks' gestation in the 1991 linked data set, including sociodemographic and medical risk factors/delivery complications.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 (i) Only linked data – birth certificate to death certificate – were used. Linkage from birth certificate to death certificate sought with 97.7% linkage achieved. (ii) Infant outcomes are based on singleton, vaginal births, 35–43 weeks' gestation (includes term births and those ± 2 weeks from term) to provide a more meaningful analysis and a suitable comparison for the CNM births. Only those risk factors which had a significant effect on birth outcomes were included in the statistical models – risk factors included were hydramnios/oligohydramnios, abruptio placentae, breech/malpresentation, fetal distress, precipitous labour, premature rupture of membranes, seizures following labour. 3 4,100,000 births used for the descriptive analysis. Of all singleton vaginal deliveries 35–43 weeks' gestation, CNM deliveries represent 5.4%, and physician deliveries 93.2%. For multivariate analysis a smaller sample totalling 810,790 births was purposively created of 153,194 certified nurse midwife deliveries (100% all CNM deliveries of singleton, vaginal births, 35–43 weeks' gestation), and a random sample of 686,644 physician deliveries (25% all physician deliveries of singleton, vaginal births, 35–43 weeks gestation) 4 Period between birth and age 1 year 5 National linked birth/infant death data set for 1991 using birth and death certificates of babies born in that period. Risk factors came from the 1989 version of the US birth certificate. Data on birth attendant came from the 'Attendant's name and title' item on the birth certificate – designed to show who actually delivered the baby.
Results Quantitative results	**Infant outcomes** *Comparison of CNM care in relation to physician care – odds ratio (95% CI) unless otherwise stated** Infant mortality: 0.81 (0.68–0.96) Neonatal mortality: 0.67 (0.48–0.94) Post-neonatal mortality: 0.86 (0.71–1.05) Low birthweight: 0.69 (0.65–0.73) Mean birthweight (OLS regression): 36.57 (p <0.01) *Adjusted for medical risk factors, complications, sociodemographic variables, gestational age and approximate duration of prenatal care

	Summary When certain sociodemographic, medical risk factors and complications are adjusted for: the risk of infant mortality was 19% lower for CNMs than for physicians; the risk of neonatal mortality was 33% lower for CNMs than for physicians; the risk of delivering a low-birthweight infant was 31% lower for CNMs than for physicians; the mean birthweight was 37 grams higher for CNMs than for physicians; and differences in postneonatal mortality were not statistically significant. Certified nurse midwives have excellent birth outcomes and provide a safe and viable alternative to maternity care in the USA, particularly for low-risk women.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjustment was made for medical risk factors and delivery complications using logistic regression: abruptio placentae; breech/other malpresentation; fetal distress; hydramnios/oligohydramnios; precipitous labour; premature rupture of membranes; seizures in labour. 2 Adjustment made for sociodemographic variables, and restriction of multivariate analysis 35–43 weeks' gestation helped to minimise bias resulting from high rates of patient transfer at gestational ages remote from term, as well as to provide comparisons. 3 Gap in data collection for two states for the education covariant, and gap in data collection for four states for smoking covariant reported. 4 No loss to follow-up reported or missing records reported. Reported gaps in data presented elsewhere in this abstract. 5 Deliberate sampling of 100% all CNM deliveries of singleton, vaginal births, 35–43 weeks' gestation and random sampling of 25% physician deliveries of singleton, vaginal births, 35–43 weeks gestation . 6 Pan-USA national study
Commentary	Limitations include gaps in data for two covariants. The person who provided prenatal care may or may not be the person delivering the baby although often they are, and also under-reporting of attendant at birth for midwives can lead to erroneous allocation of delivery to professional other than midwife. Some discrepancy in the total number of deliveries stated as analysed and the number of births reported in the study tables. It could be presumed that this difference is due to multiple births. Authors say doubt exists over data accuracy of specific variables, e.g. gestational age. Sociodemographic factors, medical risk factors and delivery complications are well reported. This is an important study that can be used to help provide evidence for the effectiveness of substitution of midwives for doctors in low-risk deliveries. However, it opens up many issues about the safety of employing large administrative data sets in this way to answer important epidemiological and health policy questions.
Research implications	A large prospective study – a controlled trial or cohort study, examining the deliveries under CNMs and physicians while recording these health outcomes across many sites would help confirm the results. Randomised controlled trials are needed to help set up care protocols that deliver the optimum health outcomes for mother and infant.

ID, origin, authors (year)	1177, UK, MacVicar, J. *et al.* (1993)
Aims	To compare the outcome of midwife led care with consultant led care during the antenatal period but especially at delivery *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwife-managed care with obstetrician-led care *Intervention/comparison:* The intervention was midwife-led care given in the antenatal period and delivery, with delivery taking place in a simulated home environment in the hospital – the Home-from-Home scheme. This was compared to consultant-led care for delivery taking place in the hospital delivery suite under a clinical environment. For women in the midwife group, mothers were looked after entirely by 10 midwives in the antenatal clinic and hospital unless consultant advice was sought. For women in the consultant group care received was hospital antenatal care under an obstetrician, shared by the general practitioner and community midwife. *Outcomes*: *Maternal* Labour outcomes: Onset of labour; raised blood pressure; intrapartum bleeding; meconium staining; electronic fetal monitoring; fetal heart irregularity; duration of labour; delay in labour; analgesia Delivery outcomes: mode of delivery; state of perineum; primary postpartum haemorrhage (PPH); manual removal of the placenta; secondary PPH; blood transfusion. Satisfaction with care: antenatal care; hospital care *Infant* Paediatrician required; birthweight; no. born <2.5 kg; no. born <37 week; Apgar score *Birth outcomes*
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial and questionnaire survey – 2: 1 midwife group to obstetrician group, taking into account the expected transfer rate from the midwife group to the obstetrician group. Zelen's method was used to randomise the consenting procedure. Staff in the control group were blinded to controls. 2 Exclusion occurred after randomisation. Exclusion criteria included: previous Caesarean section or difficult vaginal delivery; complicating general medical condition (diabetes, epilepsy, renal disease, etc.); previous stillbirth or neonatal death; previous small for gestational age baby; multiple pregnancy; Rhesus antibodies; raised serum alpha-feto protein on two occasions. Women who had had a termination, stillbirth or neonatal death were excluded from the satisfaction survey. 3 2000 subjects in total were calculated as having sufficient power to detect changes in outcome measures. 3510 women were randomised: 2304 were allocated to the Home-from-Home group and 189 refused. After consenting, 2115 remained in the experimental group and 1206 were randomised to the control group. Transfers: 45% were transferred from Home-from-Home care to specialist care and 46% those randomised were delivered in the midwife unit. 72% (2489) returned the questionnaire, and response rates from the midwife group were 73% and 69% from the obstetrician group 4 From booking until delivery and into the postnatal period 5 Randomisation occurred from 1 March 1989 until 6 July 1990, and all pregnancies were complete by February 1991. A satisfaction questionnaire was sent to women in the postnatal period 6 weeks after delivery.

Results Quantitative results	**Maternal outcomes** *Labour outcomes: midwife vs. shared care – % (*p-*value) unless otherwise stated* Onset of labour (p <0.0001): spontaneous 73 vs. 64; induced 9 vs. 11; augmented 12 vs. 16 Raised blood pressure: 4 vs. 4 (p = 0.65) Intrapartum bleeding: 2 vs. 1 (p = 0.21) Meconium staining: 15 vs. 15 (p = 0.82) Electronic fetal monitoring 50 vs. 89 (p <0.0001) Fetal heart irregularity: 22 vs. 31 (p <0.001) Duration of labour (p <0.001): duration of 1st stage (minutes) 385 vs. 355; duration of 2nd stage (minutes) 22 vs. 23; delay in 1st stage 12 vs. 12 (p = 0.82); delay in 2nd stage 8 vs 9 (p = 0.41) Analgesia (p <0.0001) not including Caesarean sections: no analgesia 13 vs. 12; nitrous oxide and oxygen 32 vs 23; pethidine or meptazinol 39 vs. 45; epidural 16 vs. 20 *Delivery outcomes: midwife vs. shared care – % (*p-*value) unless otherwise stated* Mode of delivery (p = 0.286): spontaneous vaginal 84 vs. 82; forceps or ventouse 8 vs. 10; vaginal breech 1 vs. 1; Caesarean section 7 vs. 7 State of perineum (p < 0.0001): intact perineum 33 vs. 30; episiotomy 23 vs. 31; vaginal and perineal tears 45 vs. 40 Primary PPH: 6 vs .6 (p = 0.77) Manual removal of the placenta 2 vs. 1 (p = 0.21) Secondary PPH: 1 vs. 1 (p = 0.18) Blood transfusion– 1 vs. 2 (p = 0.43) *Satisfaction with care: midwife vs. shared care – % (*p-*value) unless otherwise stated* Antenatal care: very satisfied 52 vs. 44 (CI 4.1% to 12.5%); fairly satisfied 42 vs. 47; neither 3 vs. 5; fairly dissatisfied 3 vs. 2; very dissatisfied 0 vs. 1 Hospital care: very satisfied 73 vs. 60 (CI 9.1% to 16.8%*;* fairly satisfied 21 vs. 31; neither 3 vs. 4; fairly dissatisfied 2 vs. 3; very dissatisfied 1 vs 2 **Infant outcomes** *Infant outcomes: midwife vs. shared care – % (*p-*value) unless otherwise stated* Paediatrician required: 23 vs. 25 (p = 0.17) Birthweight, mean: 3337 vs. 3348 (p = 0.58) No. born <2.5 kg: 5 vs. 5 (p = 0.87) No. born <37 weeks: 5 vs. 6 (p = 0.15) Apgar score 1 minute median: 7 vs. 8 (p = 0.11) Apgar score 5 minutes median: 9 vs. 9 (p = 0.11) Birth outcomes (p = 0.32): discharged alive and well 98 vs. 98; retained in neonatal unit 1 vs. 2; stillbirths 0 vs. 0; early neonatal deaths 1 vs. 1

	Summary For low-risk women, care provided by midwives was as safe and effective as shared care provided under an obstetrician. In the midwife group, there were more women with spontaneous onset of labour, and fewer women undergoing electronic fetal monitoring, with fewer diagnoses of fetal heart irregularity. Women in the midwife group also had a longer stage of labour (due to the lower incidence of induction and augmentation), and were more likely to use no anaesthesia or only nitrous oxide and oxygen. Women in the midwife group were more likely to have an intact perineum with fewer episiotomies but a greater number of perineal tears. There were no statistical or clinical differences in the outcomes for maternal and fetal mortality and morbidity. The midwife group were more satisfied with antenatal care and were more satisfied with care during delivery.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported. 2 Not reported. 3 No gaps in data collection reported. 4 Follow-up was by intention to treat. 5 Assignment to groups was through a sealed envelope attached to the notes containing the allocation. 6 One university teaching hospital obstetric unit
Commentary	All the staff in the midwifery group were volunteers – this may produce a bias. These staff are more likely to be dedicated, skilled and self-reliant. Demographics were reported and the groups were similar except for smoking, which was more prevalent in the control group. The midwife group also sought to provide a degree of continuity of care.
Research implications	This is the first study that has been done on a Home-from-Home midwife unit. Assessments of birth centres are few and it may be interesting to join up different midwife-led units in a multi-site study of maternal and fetal outcomes across the country. The home environment could be expected to create a more satisfying environment, but the extent to which the increased satisfaction improves the physical health outcomes is not clear, and could benefit from further investigation

ID, origin, authors (year)	480, USA, Mayes, F. *et al.* (1987)
Aims	To look for differences in care practices and outcomes between certified nurse midwife (CNM) and physician management of low-risk pregnancies *Workforce:* Tertiary care setting Nursing workforce: midwife *Feature:* Substitution of midwives for physicians *Intervention/comparison:* A pilot study comparing CNM care of pregnant women to physician care at a single study site. Unclear whether physician provider was staff, resident, or family care physician. The CNM service had been newly established. *Outcomes:* *Maternal* Medication used: • labour: analgesia/sedation; no analgesia/sedation • delivery: no anaesthetic; local or pudendal; epidural; general Perineal condition: • intact • not intact: 1st-degree perineal tear; 1st- or 2nd-degree vaginal or labial tear; 2nd-degree perineal laceration; midline episiotomy and 3rd-degree extension; midline episiotomy and 4th-degree extension; periurethral tear; spontaneous 3rd-degree laceration; spontaneous 4th-degree laceration Other various outcomes as reported
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort with matching of patient characteristics 2 Vaginal births attended by CNMs during April, May and June of 1985. All mothers were low-risk status and Caesarean sections were not included. Certain criteria (not stated) were required to be followed by the clients to qualify for this grouping and the same criteria were applied to the physician group. Cases were matched for time of delivery, exact parity, mother's age (± 5years), infant weight (within 1 lb). Most matches were found from the records as having occurred within 24–48 hrs delivery of the CNM group. 3 58 subjects in total, 29 women in each study group 4 Unclear – likely delivery and immediate postpartum period only 5 Unclear exactly – data gathered on women who delivered in spring of 1985; information was taken from maternity charts using a data reporting form.

Results Quantitative results	**Maternal outcomes** *Medication used: CNM cases vs. physician cases %* Labour: analgesia/sedation 10 vs. 45; no analgesia/sedation 90 vs. 55 (p <0.02) Delivery: no anaesthetic 55 vs. 7; local or pudendal 45 vs. 66; epidural 0 vs. 24; general 0 vs. 3 (p < 0.01) *Perineal condition: CNM cases vs. physician cases %* Intact: 28 vs. 7 Not intact: 1st-degree perineal tear 10 vs. 10; 1st- or 2nd-degree vaginal or labial tear 38 vs. 3; 2nd-degree perineal laceration 0 vs. 0; midline episiotomy and 3rd-degree extension 7 vs. 10; midline episiotomy and 4th-degree extension 0 vs. 10; periurethral tear 0 vs. 14; spontaneous 3rd-degree laceration 0 vs. 0; spontaneous 4th-degree laceration 0 vs. 0 (p <0.01) Other various outcomes as reported IV fluids given to 38% CNM group, and 72% physician group (n=58, p <0.02) Drugs to induce or augment labour given to 22% CNM group, and 56% physician group (n=54, p <0.01) Artificial rupture of membranes performed in 35% CNM group, and 66% physician group (n=58, p <0.05) Electronic fetal monitoring used in 34% CNM group, and 100% physician group (n=58, p <0.01) Birth in delivery suite (not a labour room) happened for 30% CNM group, and 89% physician group (n=54, p <0.01) Average length of 2nd-stage labour was 54.2 minutes for CNM group, and 60.2 minutes for physician group (n=58, p = 0.05) Average length of 3rd-stage labour was 16.9 minutes for CNM group, and 6.8 minutes for physician group (n=5, p <0.01) No significant differences in infant prematurity, postmaturity, baby complications, meconium, or Apgar scores **Summary** Authors do not conclude on the results that midwife care may be more or less desirable, but say: 'women will experience very different types of care according to the care provider and a larger sample size would increase the explanatory power of the findings.'
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported. 2 None reported. All appropriate tests were used on only those pairs for which data was available for both CNM and physician members. 3 Not reported. 4 None 5 None 6 One university hospital
Commentary	The authors report several important limiting factors. Only basic demographics were reported. Women were not randomly assigned to providers but self-selected themselves for nurse–midwife or physician care. This difference may have manifested itself in terms of the marital and socioeconomic statistics of the demographic factors, which varied widely. Smoking also varied between the groups, and a tendency for health-promoting behaviours among the CNM group and level of socioeconomic status may also have confounded the results. Sample size rather limited, n=29 in each group. This study is out of date in comparison to more recent work in this area and many of the research questions that this study may present are most likely answered in current research literature.
Research implications	Research in this area needs studies of robust case–control design using whole data sets, or prospective cohort studies encompassing a number of sites.

ID, origin, authors (year)	298, USA, Mehl-Madrona, L. and Mehl Madrona, M. (1997)
Aims	To study the safety and risks of attending breech, twins, and post-dates pregnancies at home for both midwives and physicians *Workforce:* Community setting – home Nursing workforce: midwife *Feature:* Substitution of midwives for physicians *Intervention/comparison:* A comparison of the outcomes of delivery of higher-risk pregnancies at home for midwife-attended births and physician-attended births. Apprentice-trained midwives are midwives who do not have a formal midwifery education and who practise outside the definition of a midwife as set by the International Confederation of Midwives (ICM). Direct-entry midwives do have formal midwifery education and are recognised by the ICM. Both types of midwife contributed to the data on which this study was based, as well as certified nurse–midwives. *Outcomes:* *Fetal* Fetal deaths before labour; fetal deaths during labour Neonatal resuscitations Neonatal deaths Total mortality Outcomes sequentially adjusted for babies with lethal congenital abnormalities, twins, breeches and post-dates Cause of death for neonates by care provider was also shown but not reported.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study with matching for patient characteristics 2 Midwife- or physician-attended births of the risk groups – twins, breech presentation and post-dates where delivered at home 3 8468 births in total, 1000 matched patients in each group drawn up. Matching was done for maternal age group, insurance status indicating socioeconomic status, parity and medical risk. 4 Retrospective study of charts of the delivery period. Data spanned years 1969–1985. 5 1970–1974: 287 midwife births, data collected by chart review by researchers. Data collected from a convenience sample 1969–1975: 355 midwife births and 791 physician births, data collected by chart review by researchers. Data collected from a convenience sample 1970–1976: 816 midwife births and 1514 physician births, data collected by chart review by researchers. Data collected from a convenience sample 1977–1985: 2593 physician births, data collected by chart review by researchers. Data offered by physicians – % cases offered by physician sample unclear 1970–1985: 3545 midwife births, self-report data collection form by midwives. Data offered by midwives – % cases offered by midwife sample unclear – 24 gave data on 1919 births in years 1970–1980; 11 gave data on 1234 births in years 1975–1985; 11 gave data on 392 births in years 1977–1983.

Results Quantitative results	**Fetal outcomes** *Fetal outcomes in matched set: no. midwife births vs. no. physician births (p-value where significant)* Entire matched set: no. 1000 vs. 1000: Fetal deaths before labour 3 vs 1; fetal deaths during labour 6 vs. 2; neonatal resuscitations 22 vs. 6 (p <0.05); neonatal deaths 5 vs. 2; total mortality 14 vs. 5 (p < 0.05) Adjustment for lethal congenital abnormalities: no. 998 vs. 997: Fetal deaths before labour 2 vs. 0; fetal deaths during labour 6 vs. 1; neonatal resuscitations 13 vs. 4 (p <0.05); neonatal deaths 4 vs. 1; total mortality: 12 vs. 3 (p <0.05) Adjustment for twins and lethal abnormalities: no. 990 vs. 996: Fetal deaths before labour 2 vs. 0; fetal deaths during labour 4 vs. 1; neonatal resuscitations 13 vs. 4 (p <0.05); neonatal deaths 4 vs. 1; total mortality 10 vs. 2 (p <0.05) Adjustment for breeches and lethal abnormalities: no. 971 vs. 994: Fetal deaths before labour 2 vs. 0; fetal deaths during labour 5 vs. 1; neonatal resuscitations 8 vs. 4; neonatal deaths 3 vs. 1; total mortality 10 vs. 2 (p < 0.05) Adjustment for post-dates and lethal abnormalities: no. 974 vs. 994: Fetal deaths before labour 0 vs. 0; fetal deaths during labour 3 vs. 1; neonatal resuscitations 9 vs. 4; neonatal deaths 3 vs. 1; total mortality 7 vs. 2 Adjustment for post-dates, breeches, twins and lethal abnormalities: no. 935 vs. 988: Fetal deaths before labour 0 vs. 0; fetal deaths during labour 1 vs. 1; neonatal resuscitations 4 vs. 4; neonatal deaths 2 vs. 1; total mortality 3 vs. 2 **Summary** There were significant differences between groups for numbers of breech deliveries, twin deliveries and post-dates deliveries that occurred at home with midwives attending more of these births. Midwives had significantly higher rates of intrapartum death and deaths before labour. No differences were found in neonatal death, but the midwife group had significantly more neonatal resuscitations. The total mortality rate was greater for midwives before and after adjusting for congenital anomalies. After adjusting for congenital anomalies, breeches, twins and post-dates, there were no more significant differences
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 A modified Popras Scoring system was used to rate medical risk at 36 weeks. It was adjusted for the variables of interest so as to award no points for breeches, twins or post-dates. 2 The neonatal and intrapartum mortality and the incidence of neonatal resuscitations were calculated after excluding infants with congenital lethal abnormalities, and the rates were further calculated after the stepwise elimination of breech presentations, twin births, and post-dates pregnancies. 3 No gaps in data reported, although data was acknowledged imperfect. 4 Not applicable 5 Matched case–control design 6 Wisconsin and western USA

Commentary	A difficult area to research. Demographics were not reported. Generalisation is limited due to differences in data collection between physicians and midwives, and non-random selection of information on births by midwives and physicians. Also women tend to self-select for home birth and predisposing factors could influence the outcome. Rarer complications would need large numbers of births analysed before comparing management. Planned and unplanned home births can be difficult to tell apart. Varying attitudes to the practice surrounding home deliveries also colours the research field. Home births are based on a low-risk delivery as the expected outcome. The author is against the delivery of higher-risk pregnancies at home.
Research implications	The argument for allowing midwives to attend high-risk births at home such as those presented by twins, breeches, or post-dates needs more focused research. Ethical issues preclude the use of randomised controlled trials in this area; therefore larger, more thorough studies of the data need to be performed, such as a prospective cohort study across many sites. This research is needed to help indicate the wisdom of attending these higher-risk births at home without ready access to medical intervention, possibly by midwives who have not attained accreditation

ID, origin, authors (year)	269, USA, Oakley, D. *et al.* (1996)
Aims	To determine whether pregnancy outcomes differ between certified nurse–midwives (CNMs) and obstetricians when statistical adjustments take alternative explanations into account – including maternal, prenatal, and intrapartum medical problems, medical processes of care, and maternal preferences *Workforce:* Tertiary care setting Nursing workforce: midwife *Feature:* Substitution of nurse–midwives for obstetricians Intervention /comparison: A study comparing the pregnancy outcomes of a group of women cared for by a group practice of CNMs during pregnancy, labour and the postpartum period, and those of a group cared for by private obstetricians. The intervention was provided by 8 CNMs and 22 obstetricians and the same facilities were used to deliver all births. *Outcomes:* *Maternal* Maternal complications: Postpartum haemorrhage; major perineal laceration; infection; medical complication; delayed bleeding; re-admission within 6 weeks; anaesthesia complications; spinal headache; severe problem; postpartum anaemia; respiratory complications; neurological complications; phlebitis *Infant* Infant outcomes: Gestational age; haematocrit; length; head circumference; 1-minute Apgar score <7; 5-minute Apgar score <7; infant bruised; respiratory difficulty; abrasions; slow, lethargic; anything abnormal; eye problem; fontanelle problem; cclavicle problem; birthweight; breast-fed at delivery; stayed with mother Transfer: to neonatal ICU; to moderate care nursery; to observation or newborn nursery Hospital charges were also reported, but not abstracted.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non randomised controlled trial 2 Exclusion criteria: hypertension needing medication during pregnancy; serious cardiac disease; chronic renal or lung disease; drug addiction; current alcoholism; seizure disorder needing medication; psychiatric illness needing medication; known multiple gestation; planned Caesarean delivery 3 1464 women recruited, 891 to the obstetrician group, 573 to the nurse–midwifery group 4 From booking until 2 months after birth for mothers, and up to 2 weeks after birth for infants 5 Data gathered between May 1988 and April 1992. 4 questionnaires were completed: (i) at booking; (ii) 32 weeks; (iii) immediately postpartum (satisfaction with care); (iv) 6 weeks postpartum. Medical charts provided the antenatal and intrapartum care information and medical outcomes. Costs were also recorded as charges.

Results Quantitative results	**Maternal outcomes** *Maternal complications by provider: nurse–midwife (n=471) vs. obstetrician (n=710) % (p-value), unless otherwise stated (NS = not significant)* Postpartum haemorrhage 14.2 vs. 25.2 (p <0.001); major perineal laceration 6.6 vs. 23.3 (p < 0.001); infection 3.6 vs. 5.9 (NS); medical complication 4.2 vs. 4.5 (NS); delayed bleeding 4.7 vs. 3.9 (NS); re-admission within 6 weeks 1.9 vs. 3.4 (NS); anaesthesia complications 0.4 vs. 1.4 (NS); spinal headache 0.4 vs. 0.3 (NS); severe problem 0.4 vs. 0.3 (NS); postpartum anaemia 0.8 vs. 0.6 (NS); respiratory complications 0.2 vs. 0.6 (NS); neurological complications 0 vs. 0.3 (NS); phlebitis 0.2 vs. 0.3 (NS) Average no. complications: nurse–midwifes: 0.37 ± 0.7; obstetricians 0.67 ± 0.88 (p < 0.001) **Infant outcomes** *Infant outcomes by provider: nurse–midwife (n=471) vs. obstetrician (n=710) % (p-value), unless otherwise stated* Gestational age (weeks) 39.45 vs. 39.42 (NS); haematocrit (%) 57.45 vs. 56.78 (NS); length (cm) 51.8 vs. 51.6 (NS); head circumference (cm) 35.04 vs. 34.82 (NS); 1-minute Apgar score <7 15.4 vs. 14.0 (NS); 5-minute Apgar score <7 2.3 vs. 2.3 (NS); infant bruised 18.3 vs. 21.5 (NS); respiratory difficulty 12.8 vs. 12.4 (NS); abrasions 3.6 vs. 6.9 (p = 0.4); slow, lethargic 5.8 vs. 5.0 (NS); anything abnormal 5.8 vs. 3.1 (p = 0.2); eye problem– 3.8 vs. 2.8 (NS); fontanelle problem 0.9 vs. 1.0 (NS); clavicle problem 0.4 vs. 0.4 (NS) Birthweight (g): <2500 3.0 vs. 2.8 (p = 0.3); 2500–4000 77.9 vs. 84.2 (p = 0.3); >4000 18.7 vs. 13.1 (p = 0.3) Breast-fed at delivery 82.2 vs. 91.7 (p < 0.001); stayed with mother 26.3 vs. 14.2 (p < 0.001) Transfer: to neonatal ICU 6.4 vs. 4.6; to moderate care nursery 6.6 vs. 5.3; to observation or newborn nursery 58.4 vs. 73.1 In obstetrician group: one stillbirth at 20 weeks and one neonatal death at 6 days due to chromosomal abnormality **Summary** The study found no significant differences between the groups for most of the indicators. The infant outcomes were excellent for both groups with no significant differences. More babies born to the midwife group were breast-fed immediately after delivery, and neonates in this group were more likely to stay with their mothers throughout her hospital stay. The midwife group experienced less haemorrhaging postpartum and were less likely to have a major perineal laceration. Women in the midwife group were also less likely to have a complication.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Regression models were used to adjust for case mix. Women's prenatal and intrapartum medical condition and baseline characteristics were controlled for by analysis of the data provided by the Problem Oriented Risk Assessment System – used to measure actual or potential medical problems and prenatal indicators conditions were variably weighted. 2 Medical processes of care and choice of provider were also controlled for. 3 For choice of provider, incomplete data collection was adjusted for by a separate regression model, testing whether choice of provider would change the conclusions. 4 Analysis was by intention to treat. 1181 remained in the study, 710 in the obstetrician group, 471 in the nurse–midwife group. Reasons for dropping out did not differ by choice of care provider. Of those who dropped out: 47% had spontaneous abortions or fetal deaths before labour; 24% could not be located, their records showing one visit only; 19% moved; 5% delivered at home or in another hospital system; 3% withdrew from the study; 2% had induced abortions. 5 None. Women self-selected for choice of provider. 6 One US mid-western hospital, its ambulatory care satellite and hospital clinics

Commentary	It is difficult to know whether the differences that remain after considering alternative explanations can be attributed to type of provider or whether there a bias in the results that could be removed with random assignment. The authors say that the findings suggest that improving outcomes will depend on reducing the pregnancy and intrapartum risks, reducing the medical processes of care, and reducing women's preference for the more expensive interventions. Women were able to choose their care provider and this is appropriate, yet maternal choice may promote concentrations of maternal characteristics. This factor is however included in the analysis. Demographics were reported and the groups were similar.
Research implications	Further research is needed to detect whether the differences that remained would be removed with random assignment, although if choice were removed, studies using random assignment may attract an unusual group of women. Analysis could be improved by a matched case–control design, yet this may be impracticable. A controlled trial therefore could be conducted to improve the grade of evidence and to help establish the efficacy of midwife care in this setting. Large samples would be needed at recruitment, to allow for attrition and statistical controls

ID, origin, authors (year)	443, USA, Olivo, L. *et al.* (1994)
Aims	To compare maternal and infant outcome variables and patient satisfaction with obstetric care provided by certified nurse–midwives (CNMs) and physicians *Workforce:* Secondary care group practice setting Nursing workforce: midwife *Feature:* Substitution of midwives for physicians *Intervention/comparison:* The satisfaction with care provided by CNMs was compared to the satisfaction with care provided by physicians as measured by the Care Provider Maternal Satisfaction Survey (CPMSS). *Outcomes:* *Maternal* Satisfaction: Satisfied with care provider; provider answered questions; good obstetric experience; would use provider again; aware of types of providers Delivery: primary Caesarean; previous Caesarean; repeat Caesarean; vaginal birth after Caesarean (VBAC) *Infant* Infant birthweight
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study and questionnaire survey 2 A convenience sample of (private) patients choosing physician care or midwifery care – source of data not otherwise reported. 3 535 patients in total. Midwife care n=225; physician care n=310. Total response rate to questionnaire n=461: midwife care n=285 (91.9%); physician care n=176 (78.2%) 4 Cross-sectional questionnaire survey covering the period between antenatal booking and post-natal discharge 5 Data collected with aid of Care Provider Maternal Satisfaction Survey (CPMSS) in 1988 and 1989. The questionnaire enquired about early prenatal care; late prenatal care; and postpartum care. A copy of survey questionnaire was provided.
Results Quantitative results	**Maternal outcomes** *Satisfaction: midwives vs. physicians % (p-value)* Satisfied with care provider: 96.8 vs. 90.9 (p <0.05) Provider answered questions: 96.5 vs. 93.8 (p >0.05) Good obstetric experience: 97.2 vs. 90.3 (p <0.05) Would use provider again: 89.9 vs. 82.4 (p <0.05) Aware of types of providers: 96.1 vs. 84.1 (p <0.0001) *Delivery: midwives vs. physicians % (p-value)* Primary Caesarean: 9.3 vs. 12.0 (p > 0.05) Previous Caesarean: 3.2 vs. 7.1 (p > 0.05) Repeat Caesarean: 70 vs. 62.5 (p > 0.05) VBAC: 30 vs. 37.5 (p >0.05) **Infant outcomes** *Infant birthweight: midwives vs. physicians g (± SD)* Birthweight: 3435 (± 663) vs. 3357 (± 779)

	Summary The midwife group were significantly more satisfied and were significantly more likely to plan for using the same care provider for future needs and were significantly more satisfied with the level of experience of the care provider. The midwife group were more informed about the availability of different types of providers. High levels of satisfaction were expressed by all women, but the midwife group were significantly more satisfied and scored more highly on 4 of 5 satisfaction measures. No significant differences were found in delivery outcomes as measured by Caesarean section rates, or successful vaginal birth after delivery.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported. 2 None reported. 3 No gaps in data collection reported. 4 Not specified. No transfer between groups reported. 5 None reported. 6 One group practice in New York, with delivery taking place at a university hospital
Commentary	No attempt made at case mix – basic demographic variables only shown and the groups were significantly different for parity, and were dissimilar in education. No indication if patients were all low-risk births and small samples used. Small sample size limits validity of the results. Patient satisfaction surveys are also prone to measurement errors. Overall, the study could have been reported better. Study also includes a review element. The findings of the study suggest that patients are satisfied with care from CNMs, and can expect similar obstetric outcomes to those found under physician care.
Research implications	Non-random sampling limits the use of these results and well-designed prospective studies or randomised controlled trials with samples of greater size and with clear inclusion and exclusion criteria that are better able to show significant differences in care are needed to help establish the efficacy of midwife-managed care plans. Protocols for midwife care that allow low-risk women to be cared for safely need developing and testing through pilot studies and further research.

ID, origin, authors (year)	1179, UK, Renfrew, M.J., Cochrane Pregnancy and Childbirth Database (1995) Issue 1; date of last substantive amendment 1992
Aims	To compare effects of midwife versus medical/shared care in pregnancy, labour, delivery and postpartum, on perinatal mortality and physical and psychosocial measures of maternal and infant morbidity *Workforce:* Midwife; Secondary *Feature:* Substitution *Outcome:* Neonatal resuscitation, admission to special care nursery, induction of labour, feelings of dissatisfaction with pain relief, the use of and the amount of pharmacological analgesia used, incidence of Caesarean section, Apgar score of less than 8 at one minute, stillbirth and neonatal death
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	 1 Early systematic review. 2 Controlled comparisons of care provided by midwives or nurse–midwives with access to medical back-up versus care provided by doctors during pregnancy, labour, delivery and postpartum 3 Individual study design: controlled trials. 4 3 studies; total number of participants unknown 5 Unknown 6 Unknown 7 Unknown
Results Quantitative results	Results of meta-analysis (pooled log odds ratio) Care given by midwives was associated with a reduction in the occurrence of: clinic waiting time of 15 minutes or more; poor ability to discuss anxieties in pregnancy or problems postpartum; feeling ill-prepared for labour; lack of enjoyment; not feeling in control during labour; feeling ill-prepared for child care; augmentation of labour; regional anaesthesia/analgesia; episiotomy; operative vaginal delivery; birthweight <2500 g; neonatal resuscitation; or admission to special care nursery. There were no differences in: induction of labour; feelings of dissatisfaction with pain relief; the use of and the amount of pharmacological analgesia used; incidence of Caesarean section; Apgar score of less than 8 at one minute; stillbirth; or neonatal death.
Commentary	A pre-Cochrane review. The reportage is incomplete by current Cochrane standards; therefore it is difficult to judge how conclusive this review could be. The author points out that many of the results given in the review are based on only one trial. The detail and summaries provided of the included studies are very slight. One reviewer compiled the review.
Research implications	A thorough systematic review is needed to expand this work. Further trials in populations that are not low risk are needed to help establish the extent of the professional boundaries between midwifery and obstetrical care.

ID, origin, authors (year)	196, UK, Spurgeon, P. *et al.* (2001)
Aims	To look at maternal satisfaction with two midwife pilot schemes based on the Changing Childbirth initiative and to compare these to a traditional model of care *Workforce:* Secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwives of doctors; team work amongst midwives *Intervention/comparison:* Maternal satisfaction in experimental groups A and B receiving care of one-to-one midwifery-led provision, compared to a third group C receiving standard obstetric-led care. The pilot groups A and B consisted of women from practices in the Changing Childbirth scheme. Women in group A were cared for by one of five named midwives, while women in group B were cared for by one of five midwives working in a team. Group C women were referred to the hospital in the normal way – receiving shared care between GP and hospital, outside of the Changing Childbirth scheme. The Changing Childbirth report (1993) advocated a woman-focused readily accessible, responsive and effective service in which women were involved in planning of the service. Providers of maternity care were given guidelines for action – to help afford (a) choice, (b) control and (c) continuity of care for the woman. Women should: have sufficient information; choose the place of birth; choose the type of care received; choose which professionals provide it; have the entitlement of a named midwife or lead professional to help them develop a birth plan and to facilitate continuity of care. There should be development of trust between client and care-giver. Greater responsibility for care to be given to midwives and GPs. *Outcomes*: *Maternal* Outcomes of labour and delivery: mean length of labour; use of pain control; normal delivery; instrumental delivery; elective Caesarean; emergency Caesarean; perineal tearing; episiotomy *Satisfaction outcomes* Personal preferences; antenatal care; labour and delivery; postnatal care; information and advice ***Infant*** Apgar score; birthweight
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study and questionnaire survey 2 No specific inclusion or exclusion criteria stated. Low-risk women only were studied. Groups drawn from practices who followed standard protocols and those who followed the Changing Childbirth scheme. 3 Group A n=112 selected from 4 GP practices; group B n=103 selected from 3 GP practices; group C n=118 selected from similar practices; total = 333 4 Over 6 weeks following delivery 5 Questionnaire for the three groups A, B and C focused on the antenatal period, delivery and the postnatal period. Data collected over 18 months in 1997–1998. Questionnaire received by women 6 weeks after birthing.

Results Quantitative results	**Maternal outcomes** *Labour and delivery: A midwives vs. B midwives vs. C shared care % (unless otherwise stated)* Mean length of labour 6 hours 8.7 minutes vs. 7 hours 17.4 minutes vs. 6 hours 26.8 minutes Use of pain control: • Yes 92.9 vs. 90.3 vs. 87.3 • No 7.1 vs. 9.7 vs. 12.7 Normal delivery 74.1 vs. 68 vs. 69.5; instrumental delivery 13.4 vs. 11.7 vs. 12.7; elective cCaesarean 7.1 vs. 7.8 vs. 5.1; emergency Caesarean 13.4 vs. 12.6 vs. 12.7 Perineal tearing: • Yes 44.6 vs. 36.9 vs. 49.2 • No 55.4 vs. 63.1 vs. 50.8 Episiotomy: • Yes 33.9 vs. 17.5 vs. 33.9 • No 66.1 vs. 82.5 vs. 66.1 *Personal preferences of health professional to manage care: A midwives vs. B midwives vs. C shared care % (unless otherwise stated)* No significant differences were found between the groups in terms of choice of venue for delivery Significant differences were found for Groups A and B: • Had more choice about where they could give birth 67 vs. 62 vs. 40: χ^2 = 19.22 (p <0.001) • Had more choice about who would deliver the baby 64 vs. 63 vs. 25: χ^2 = 45.14 (p <0.001) • Believed that they had a contact when advice or information was needed 98 vs. 91 vs. 80: χ^2 = 24.15 (p <0.001) Group A rated the value of the antenatal/parentcraft classes more highly than the other two groups: χ^2 = 13.24 (p <0.05) *Antenatal care* Significant differences were found for the following: • Group C felt they had too few checks at home compared to the other groups: χ^2 = 9.92 (p <0.05) • Group C had their first antenatal check at a GP clinic, compared with groups A and B whose first check-up was at home: χ^2 = 46.48 (p <0.001) For health professionals seen at first and subsequent check-ups, it seems that although the control group was likely to be offered a named midwife throughout, where this was not available, care was provided by a greater range of people. Groups A and B were significantly more satisfied than group C with information relating to choice of: • venue for delivery: χ^2 = 18.56 (p <0.01) • provision of care: χ^2 = 24.83 (p <0.001) • type of maternity care available– χ^2 = 24.79 (p <0.001) • details of care: χ^2 = 17.91 (p <0.01) • preparation for labour: χ^2 = 17.95 (p <0.01) No difference was found between the groups with regard to the amount of information provided about: choice of hospital or GP unit; where check-ups could be conducted; preparation for motherhood. Groups A and B were significantly more satisfied with: • care and sensitivity of the staff: F = 5.43 (p <0.01) • Contact with midwives: F = 17.73 (p <0.001) • sense of not being pressured: F = 4.5 (p <0.05) • mother's views being taken into account: F = 7.32 (p <0.05) • consistency of information and advice: F = 5.87 (p <0.01)

	Labour and delivery: A midwives vs. B midwives vs. C shared care % (unless otherwise stated) For knowing the midwife who delivered them: 92% vs. 94 % vs. 8% Groups A and B were more satisfied than group C with their midwives during labour and delivery with: • the degree of explanation about what was happening: $F = 8.45$ ($p < 0.001$) • kind and understanding behaviour: $F = 8.42$ ($p < 0.001$) • attention to women's needs: $F = 5.73$ ($p < 0.01$) • response to women's requests: $F = 4.15$ ($p < 0.05$) • not leaving women alone too much: $F = 5.26$ ($p < 0.01$) No difference was found between for satisfaction with pain relief. *Postnatal care: A midwives vs. B midwives vs. C shared care % (unless otherwise stated)* No significant differences in length of stay Groups A and B had more home visits by the midwife: $F = 25.71$ ($p < 0.001$) (These visits were not prompted by concerns about fetal wellbeing.) Significant differences were found with satisfaction expressed on all aspects of postnatal care received by groups A and B for: • care and sensitivity of midwife: $F = 10.07$ ($p < 0.001$) • explanation/consultation about concerns: $F = 3.64$ ($p < 0.05$) • help with feeding: $F = 6.84$ ($p < 0.05$) • monitoring baby's health/progress: $F = 3.44$ ($p < 0.01$) • monitoring mother's health/progress: $F = 5.69$ ($p < 0.01$) • taking maternal views into account: $F = 6.44$ ($p < 0.001$) • information/advice provided: $F = 11.45$ ($p < 0.001$) • willingness of midwife to attend to needs: $F = 11.64$ ($p < 0.001$) *Information and advice* No significant differences found between the groups in the adequacy of the information given about antenatal tests including: routine booking blood test; AFP test; ultrasound; additional scans; amniocentesis; CVS No significant differences found between the groups in the satisfaction with information provided before birth, in hospital, and at home including: feeding methods; the baby's health; handling; washing and changing baby; possible problems and complications; and information for fathers Significant differences were found with satisfaction expressed retrospectively on the level of information received by groups A and B for: • choice of birth: $\chi^2 = 20.11$($p < 0.01$) • pain relief: $\chi^2 = 23.3$ ($p < 0.001$) • different drugs used in labour: $\chi^2 = 23.37$ ($p < 0.001$) **Infant outcomes** *Fetal outcomes: A midwives vs. B midwives vs. C shared care* Mean Apgar score 1 minute: 8.38 vs. 8.33 vs. 8.25 Mean Apgar score 5 minutes: 8.98 vs. 9.06 vs. 0.64 Birthweight (kg: 3.36 vs. 3.24 vs. 3.29

	Summary There were no significant differences in maternal and neonatal clinical outcomes. Among women there was a general preference for and high satisfaction with continuity of midwifery-led care rather than carer, and given the high levels of satisfaction and good clinical outcomes with midwifery-led care, there is a case for making this model of care more available. Midwifery-led care was much preferred to obstetrician-led care and did not lead to any deficits in clinical outcomes. One of the two midwife pilot schemes showed no reduction in satisfaction levels or other outcome measures.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No adjustment for case mix. All women of high obstetric risk were excluded from the study. 2 Unmatched groups, but no significant differences existed between the groups in terms of personal or clinical characteristics. 3 No gaps in data collection reported. 4 No loss to follow-up reported. 5 Non-random, naturalistic allocation to groups 6 One trust in which seven practices in the Changing Childbirth scheme were involved and were from the same area. Other practices outside the scheme from the trust were used in the study. The trust crossed a range of socioeconomic strata.
Commentary	This retrospective study is potentially flawed due to limitations and distortions of participants' memories. A holistic view of continuity was considered more achievable through a retrospective design. It is also possible that the groups could rate their care highly as a function of their experience – women could overrate their experience to vindicate their continued involvement in the study. An economic analysis would have shown up the costs of extra visiting that the pilot groups received. Demographics were not well reported, only sample characteristics of age and parity, for which the groups were similar.
Research implications	There is a good case to make the midwife model of care more available. An economic analysis should be completed in parallel with a randomised controlled trial of good design to present the costs of delivering such care. Further proper, reliable and valid studies on estimating the satisfaction with care of tried, effective midwife care models could help to promote the case for more midwife responsibility of care for low-risk births

ID, origin, authors (year)	934, UK, Tucker, J.S. *et al.* (1996)
Aims	To compare the routine antenatal care of general practitioners and midwives with that of obstetrician-led shared care *Workforce:* Primary care and community setting Nursing workforce: midwife Medical workforce: general practitioner *Feature:* Substitution of joint midwife and GP care for obstetrician care Intervention/comparison: The intervention is routine antenatal care by general practitioners and midwives according to a care plan with protocols for managing complications compared to obstetrician-led shared care. *Outcomes:* Comparisons of care included clinical evaluation, measures of women's satisfaction and of staff satisfaction (not reported), and a health economic analysis (not reported). Data on Health Service use were reported but not presented in this abstract. *Maternal* Failures of care: Failure to: diagnose anaemia after blood test; treat anaemia after blood test; refer malpresentation to specialist; refer at 42 weeks' gestation to specialist; check Rhesus-negative women for antibodies Diagnosed antenatal complications in women of low risk: Pregnancy induced hypertension; transient hypertension; proteinuria; pre-eclampsia; anaemia; multiple pregnancy; malpresentation/unstable lie; antepartum haemorrhage; gestational diabetes; hydramnios; hyperemesis; urinary tract infection; other condition Intrapartum events and pregnancy outcomes: No medical notes at admission in labour Undiagnosed conditions at admission in labour: hypertension; multiple pregnancy; malpresentation; intrauterine death; other condition Labour type: spontaneous; induced; augmented; planned Caesarean Preterm delivery <37 weeks Pregnancy outcome: live birth; stillbirth; early neonatal loss; fetal loss <24 weeks; termination Mode of delivery: spontaneous vaginal; forceps or ventouse; breech vaginal; emergency Caesarean; earlier than planned Caesarean; planned elective Caesarean Undiagnosed abnormality at birth Baby in special care baby unit (SCBU) >48 hours Baby breast-fed in hospital Women's satisfaction with aspects of their care: Overall satisfaction; acceptability of style; relationship with staff; experience attending clinics; information acquisition; service access and provision Full results for satisfaction are abstracted only where differences were significant.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Pragmatic multicentre randomised controlled trial and questionnaire survey 2 Exclusion criteria employed, but not explicitly stated. Low-risk mothers <18 weeks' gestation only entered trial – those at high risk of antenatal complications were excluded (based on the presence of a previous obstetric history, previous Caesarean section, current pregnancy conditions, or serious medical conditions 3 1765 eligible women consented to join trial – 834 under GP and midwife care, 840 under obstetrician-led care 4 From time of first booking visit (<18 weeks) until delivery and through to 6 weeks postpartum 5 Data collection was over 13 months in years 1993–1994. 224 general practitioners and 45 community midwives were involved in giving care. Demographic data were collected from the record of the booking visit. Clinical data were abstracted from the medical records, shared care cards, and midwifery records after delivery. For quality of antenatal care received, data came from the hospital. Of the 1765 eligible women who consented, 1712 (97%) received a copy of the satisfaction questionnaire excluding those who had aborted, had terminations, stillbirths, neonatal death, or whose babies were in SCBU.

Results Quantitative results	**Maternal outcomes** *Failures of care: GP and midwife care vs. specialist care % (*p-*value)* Failure to: diagnose anaemia after blood test 0.3 vs. 0.2 (p = 1.0); treat anaemia after blood test 0.0 vs. 0.7 (p = 0.04); refer malpresentation to specialist 0.9 vs. 0.0 (p = 0.25); refer at 42 weeks' gestation to specialist 0.2 vs. 0.0 (p = 0.48); check Rhesus-negatve women for antibodies 2.5 vs. 0.4 (p = 0.0008) *Diagnosed antenatal complications in women of low risk: GP and midwife care vs. specialist care % (*p-*value)* Pregnancy induced hypertension 4.4 vs. 8.4 (p = 0.002); transient hypertension 8.2 vs. 11.1 (p = 0.04); proteinuria 9.6 vs. 13.9 (p = 0.007); pre-eclampsia 1 vs. 4 (0.0005); anaemia 13.6 vs. 13.1 (p = 0.8); multiple pregnancy 0.4 vs. 0.7 (p = 0.5); malpresentation/ unstable lie 4.8 vs. 3.9 (p = 0.5); antepartum haemorrhage 2.5 vs. 3.0 (p = 0.7); gestational diabetes 0.8 vs. 0.7 (p = 1.0); hydramnios 0.8 vs. 1.0 (p = 1.0); hyperemesis 0.4 vs. 1.1 (p = 0.2); urinary tract infection 8.4 vs. 7.0 (p = 0.3); other condition 12 vs. 12 (p = 0.8) *Intrapartum events and pregnancy outcomes: GP and midwife care vs. specialist care % (*p-*value)* No medical notes at admission in labour 1.8 vs. 1.1 (p = 0.3) Undiagnosed conditions at admission in labour: hypertension 0.4 vs. 0.2 (p = 0.7); multiple pregnancy– 0 vs. 0 (p = 0); malpresentation 0.9 vs. 0.2 (p = 0.2); intrauterine death 0.1 vs. 0.1 (p = 0); other condition– 0.2 vs. 0.5 (p = 0.5) Labour type: spontaneous 58.5 vs. 51.5 (p = 0.009); induced 18.1 vs. 24.5 (p = 0.009); augmented 20.1 vs. 20.9 (p = 0.009); planned Caesarean 3.2 vs. 3.0 (p = 0.009) Preterm delivery <37 weeks 5 vs. 5 (p = 0.8) Pregnancy outcome: live birth 97.8 vs. 96.8 (p = 0.5); stillbirth 0.5 vs. 0.4 (p = 0.5); early neonatal loss 0.2 vs. 0.6 (p = 0.5); fetal loss <24 weeks 1.1 vs. 1.8 (p = 0.5); termination 0.4 vs. 0.5 (p = 0.5) Mode of delivery: spontaneous vaginal 78.9 vs. 80.0 (p = 0.4); forceps or ventouse 11.9 vs. 9.7 (p = 0.4); breech vaginal 0.4 vs. 0.6 (p = 0.4); emergency Caesarean 5.9 vs. 7.4 (p = 0.4); earlier than planned Caesarean 0.2 vs. 0.2 (p = 0.4); planned elective Caesarean 2.7 vs. 2.0 (p = 0.4) Undiagnosed abnormality at birth 1.4 vs. 1.4 (p = 0.9) Baby in special care baby unit (SCBU) >48 hours 5.9 vs. 7.7 (p = 0.2) Baby breast-fed in hospital 47 vs. 48 (p = 0.6) *Women's satisfaction with aspects of their care*: *GP and midwife care vs. specialist care %* Overall satisfaction: Did you enjoy your care? (p = 0.04): yes 70 vs. 63; usually 25 vs. 31; not very much 4 vs. 5; not at all 1 vs. 1 How satisfied were you with the care you received during your pregnancy? (NS) Acceptability of style: Were you happy with the arrangement of your antenatal visits? (NS) Did you want to see a hospital doctor but didn't? (NS) Relationship with staff: How well did you get on with your main carer? (p = 0.04): very well 71 vs. 67; reasonably well 29 vs. 31; not very well 0 vs. 2 Preferred level of continuity of care (p < 0.0001): didn't mind someone different each time 13 vs. 18; small group of 3–4 people 13 vs. 15; one person but didn't mind someone different 45 vs. 49; same person each time 29 vs. 18 Experience attending clinics: Waiting times at health centre clinics (NS) Waiting times at hospital clinics (NS) Information acquisition: How satisfied are you with information about preparation for labour? (NS) Service access and provision: Did you go to antenatal classes? (NS) Did you visit the labour rooms in hospital before you came in to have your baby? (NS)

	Summary There were gains for the GP–midwife group in terms of antenatal continuity of carer, and fewer non-attendances, day care episodes and admissions in the antenatal period. Also there was a statistically significant but numerically small reduction in the number of routine clinic visits for women in the GP–midwife group, with fewer routine visits for multiparous women than primiparous women. There were few failures to care in both groups. Significantly more Rhesus–negatve women in the GP–midwife group did not have an antibody check-up. Significantly more women with anaemia on testing in the obstetrician group did not receive treatment. Fewer women developed pregnancy-induced hypertension, proteinuria or pre-eclampsia, while both groups had similar numbers of women with undiagnosed hypertension at admission in labour. Both study groups reported they were happy with care, but some differences emerged – women in the GP–midwife group reported a better relationship with their general practitioner and a stronger preference to see the same person at each antenatal visit.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 None reported. 3 No gaps in data collection reported. 4 Analysis done by intention to treat – of the 1765 eligible women who gave consent, at follow-up 9 women were withdrawn from trial but were included in follow-up. A further 91 women had incomplete medical records (44 in the GP and midwife group, and 47 in the shared care group) – 5 % in each group. The clinical evaluation of the remaining 1674 women was reported. 5 Randomisation between groups using opaque envelopes. 6 Dispersal of 51 practices linked to 9 maternity hospitals – these hospitals provide care for 38% maternity population of Scotland.
Commentary	Exclusion criteria were not explicitly stated with reports only of low-risk women entering the study, and this may make it more difficult to compare across studies. However, demographics were well reported and no significant differences existed between the groups at baseline. This study provides further evidence suggesting that antenatal care for normal women can be safely handed over to primary care professionals. The study also shows that antenatal visits at specialist clinics for these women can be made on the basis of need with no detriment to health outcomes.
Research implications	An economic analysis could be most useful in determining the financial implications for a switch to this form of care plan and the results of that research may go to inform the feasibility of a regional health policy of recommending all low-risk births be placed under the care of midwives.

ID, origin, authors (year)	935, UK, Turnbull, D. *et al.* (1996)
Aims	To compare midwife-managed care and shared care in terms of clinical efficacy and maternal satisfaction *Workforce:* Mixed primary and secondary care setting Nursing workforce: midwife *Feature:* Substitution of midwife-led care for doctor-led care *Intervention/comparison:* A comparison of midwife-only managed care on a midwife development unit with shared care between midwives, hospital doctors and general practitioners *Outcomes:* *Maternal* Relevant interventions as health outcomes: Intrapartum: induction of labour; augmentation of labour Condition of perineum: intact; episiotomy; tear Pain relief (excluding elective Caesarean section): none; tens/Entonox/bath; pethidine/piamorphine; epidural Maternal outcomes: Mean duration of labour by stage 1, 2 and 3; number at gestation when delivered (25–36 weeks; 37–41weeks; ≥42weeks); number by mode of delivery; number with manual removal of placenta Maternal complications: Antenatal: antepartum haemorrhage; anaemia; hypertension; major medical complications; multiple pregnancy; placenta praevia Intrapartum: antepartum haemorrhage; cord presentation; cord prolapse; hypertension; inverted uterus; malpresentation; postpartum haemorrhage Postnatal: hypertension; major medical complication; postnatal depression; postpartum haemorrhage Overall satisfaction with maternity care: Antenatal; intrapartum; hospital-based postnatal; home-based postnatal Overall satisfaction with maternity care presented in this abstract only. Satisfaction with choice, information, decision-making, and individualised care was also reported for the antenatal, intrapartum, hospital-based postnatal, and home-based postnatal periods, but were not presented in this abstract due to limitations on space. *Infant* Fetal and neonatal outcomes: Re-admissions; birthweight centile for gestational age; Apgar score; neonatal standby requested; pre-admission observation in special care baby unit (SCBU); admission to SCBU Neonatal complications and fetal and neonatal loss: Neonatal complication: birth asphyxia; bowel obstruction; cardiac problems; fitting/seizures; hypoxic encephalopathy; significant jaundice; meconium aspiration; pneumothorax; fetal loss before 24 weeks: spontaneous abortion; induced abortion; fetal loss after 24 weeks; stillbirth; neonatal death

Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial and survey questionnaire 2 Criteria for eligibility: residence with hospital catchment area; booking for antenatal care within 16 full weeks pregnancy; absence of medical or obstetric complications (complications not stated in study) 3 1299 women consented to participate – 648 assigned midwife care and 651 assigned shared care. Response rate to the questionnaires: third trimester questionnaire 85.3% of midwife group vs. 78.2% of shared-care group; postpartum questionnaire 71.9% of midwife group vs. 63.1% of shared-care group. 4 Follow-up time was from period of booking until 28 days postpartum for mothers, and from birth until transfer out of the place of delivery for babies. 5 Recruitment: January 1995 to February 1994. 20 midwives of the midwife development unit provided care. Clinical data: taken from maternity case record with data on women's care, including interventions, outcomes and complications; shared-care card, a liaison document; midwifery kardex used for admissions, intrapartum and hospital-based postnatal care. Additionally the MDU care plan for women in midwife-managed group. For women, records covered period from booking to 28 days postpartum. For babies, data were collected from birth until transfer to health visitor, or special care baby unit (SCBU). Satisfaction with care: to measure satisfaction with care, two self-report questionnaires gathered information on antenatal care and postpartum care and were sent to all women except those who suffered a miscarriage or stillbirth. The antenatal care was assessed after 34–35 weeks' gestation in the third trimester. The postpartum care was assessed after 7 weeks postpartum.
Results Quantitative results	**Maternal outcomes** *Relevant interventions as health outcomes: midwife care vs. shared care % (95% CI), unless otherwise stated* Intrapartum: induction of labour 23.9 vs. 33.3 (4.4 to 14.5); augmentation of labour 43.1 vs. 39.7 (–9.0 to 2.1) Condition of perineum (p = 0.02): intact 30.5 vs. 23.6; episiotomy 28.0 vs. 34.0; tear, 1st or 2nd degree 41.5 vs. 42.4 (p = 0.02) Pain relief excluding elective Caesarean section (p = 0.005): none 12.8 vs. 11.9; tns/Entonox/bath 12.0 vs. 9.0; pethidine/diamorphine 42.6 vs. 45.1; epidural– 32.7 vs. 34.1 *Delivery outcomes: midwife care vs. shared care % (95% CI), unless otherwise stated* Mean duration of labour (hours): stage 1 6.4 vs. 6.1 (–0.7 to 0.2); stage 2 1.0 vs. 1.0 (–0.1 to 0.1); stage 3 0.2 vs. 0.3 (–0.1 to 0.1) Number at gestatation when delivered (weeks): 25–36 weeks 4.9 vs. 7.0 (–0.5 to 4.8); 37–41 weeks 93.1 vs. 91.1 (–5.0 to 1.0); 42 or >42 weeks 2.0 vs. 1.8 (1.7 to 1.4) Number by mode of delivery: spontaneous vertex 73.5 vs. 73.7 (p = 0.9); instrumental 13.6 vs. 14.3 (p = 0.9); emergency section 9.8 vs. 9.2 (p = 0.9); elective Caesarean 3.1 vs. 2.7 (p = 0.9) Number with manual removal of placenta: 4.0 vs. 4.0 (–2.2 to 2.1) *Maternal complications: midwife care vs. shared care % (95% CI), unless otherwise stated* Antenatal: antepartum haemorrhage 5.4 vs. 5.5 (–2.4 to 2.5); anaemia 18.4 vs. 19.6 (–3.2 to 5.6); hypertension 4.8 vs. 10.0 (2.3 to 8.0); major medical complications 0.3 vs. 0 (–0.7 to 0.1); multiple pregnancy 0.8 vs. 0.6 (–1.1 to 0.8); placenta praevia 2.5 vs. 1.6 (–2.5 to 0.6) Intrapartum: antepartum haemorrhage 1.6 vs. 3.6 (0.2 to 3.8); cord presentation 0.2 vs. 0.0 (–0.5 to 0.2); cord prolapse– 0.2 vs. 0.0 (–0.5 to 0.2); hypertension 3.4 vs. 2.5 (–2.8 to 1.0); inverted uterus 0.2 vs. 0.0 (–0.5 to 0.2); malpresentation 4.2 vs. 2.8 (–3.5 to 0.7); postpartum haemorrhage 5.9 vs. 5.7 (–2.8 to 2.4) Postnatal: hypertension 3.7 vs. 4.9 (–1.2 to 3.4); major medical complication 0.2 vs. 0.2 (–0.4 to 0.5); postnatal depression 0.3 vs. 0.2 (–0.7 to 0.4); postpartum haemorrhage 0.3 vs. 0.3 (–0.6 to 0.7) *Overall satisfaction with maternity care: midwife care vs. shared care mean score (95% CI)* Antenatal: 1.41 vs. 0.93 (–0.55 to –0.41) Intrapartum: 1.49 vs. 1.21 (–0.37 to –0.18) Hospital-based postnatal; 1.34 vs. 0.77 (–0.70 to –0.45) Home-based postnatal: 1.45 vs. 1.11 (–0.42 to –0.25)

	Infant outcomes *Fetal and neonatal outcomes: midwife care vs. shared care % (95% CI), unless otherwise stated* Re-admissions: 2.0 vs. 1.2 (–2.3 to 0.6) Birthweight centile for gestational age: <5th centile 4.1 vs. 4.8 (–1.6 to 3.0); 5th–9th centile 88.1 vs. 88.6 (–3.1 to 4.1) Apgar score: 8–10 at 1 minute 78.2 vs. 75.7 (–7.2 to 2.2); 8–10 at 5 minutes 97.8 vs. 96.6 (–3.1 to 0.6) Neonatal standby requested: 48.6 vs. 50.1 (–4.1 to 7.1) Pre-admission observation in SCBU: at birth 19.4 vs. 17.2 (–6.5 to 2.2); postnatally 5.6 vs. 5.5 (–2.6 to 2.5) Admission to SCBU: at birth 5.4 vs. 6.6 (–1.4 to 3.9); postnatally 3.8 vs. 3.0 (–3.0 to 1.3) *Neonatal complications and fetal and neonatal loss: midwife care vs. shared care % (95% CI), unless otherwise stated* Neonatal complication: birth asphyxia 0.6 vs. 0.8 (–0.8 to 1.1); bowel obstruction 0 vs. 0.2 (–0.2 to 0.5); cardiac problems 0.6 vs. 0.7 (–0.9 to 0.9); fitting/seizures 0 vs. 0.7 (–0.2 to 0.5); hypoxic encephalopathy 0 vs. 0.2 (–0.2 to 0.5); jaundice requiring Rx 9.6 vs. 8.8 (–4.1 to 2.4); meconium aspiration 0.2 vs. 0 (–0.5 to 0.2); pneumothorax 0.2 vs. 0 (–0.5 to 0.2) Fetal loss before 24 weeks: spontaneous abortion 31 vs. 3.8 (–1.3 to 2.8); induced abortion 0.2 vs. 0.2 (–0.4 to 0.5) Fetal loss after 24 weeks: stillbirth 0.2 vs. 0.7 (–0.2 to 1.2); neonatal death 0.5 vs. 0.8 (–0.6 to 1.2) **Summary** The authors found that midwife-managed care resulted in similar or reduced rates of interventions, similar outcomes, similar complications for mother and baby, and greater satisfaction with care, supporting the original hypothesis. Overall women were satisfied with care in both groups but women in the midwife group expressed significantly greater satisfaction overall than women in the shared care group, for all stages, with the greatest differences found in satisfaction with antenatal care and hospital-based postnatal care. Women in the midwife group were more satisfied with choice, information, decision making, and individualised care. Midwife-managed care for healthy women is clinically efficacious.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case-mix adjustment reported. 2 None stated. 3 Gaps in data collection were reported. The shared-care card was available for only 82% women in midwife-managed care, as opposed to 59.3% in the shared care group and therefore analysis of outcomes based on this record was restricted to women with complete data; no significant differences were shown for the variables between those excluded and those with complete data. The case record was available for 97.5% of the shared care group and 99.2% of the midwife-managed group. 4 Analysis by intention to treat. Follow-up until 28 days postpartum, and neonatal transfer off labour ward. 5 Allocation between groups with restricted randomisation using random numbers tables 6 One maternity hospital in Glasgow catering for a deprived area
Commentary	Medical exclusion criteria not reported, therefore difficult to compare the level of maternal risk directly to other studies. Demographics are reported as socioeconomic status and parity for mothers – no significant differences were found between groups. An Apgar score was also reported for infants. Generalisability should be attempted with caution, as the intervention was delivered in an integrated maternity service in a consultant obstetric unit, which may affect practice and study outcomes. The results suggest maternity services can be delivered efficaciously by midwives for healthy women.
Research implications	More studies in other maternity units are needed to evaluate midwife-managed care in a range of settings. A multi-site randomised controlled trial with adoption of this successful midwifery development unit style of care in other regional units is needed to help establish its potential for uptake.

ID, origin, authors (year)	1174, Australia. Waldenstrom, U. *et al.* (2000) – linked to study 168
Aims	To evaluate the effect of team mjdwife care on satisfaction with antenatal, intrapartum, and postpartum care in women at low medical risk in early pregnancy *Workforce:* Secondary care antenatal clinic setting Nursing workforce: midwife *Feature:* Substitution of midwife team care for standard models of care; operation of workforce with team work among midwives *Intervention/comparison:* A team of 8 midwives providing antenatal, intrapartum and postnatal care in collaboration with medical staff using the same medical protocols as standard care. Standard models of care included: (i) care from doctors mostly in a hospital clinic; (ii) care from midwives mostly in a midwives clinic (in collaboration with hospital doctors); (iii) birth centre care mostly run by midwives; (iv) shared care between general practitioners and hospital doctors. *Outcomes:* *Maternal outcomes* Satisfaction with antenatal care; satisfaction with intrapartum care; satisfaction with postnatal care *Process of care outcomes* Number of antenatal visits by health professional; number of antenatal caregivers by type; waiting time at antenatal visits; attended childbirth education classes; number of classes; number of midwives during labour and birth; seen in labour by midwife, seen antenatally; team midwife present; accoucher; seen postnatally by midwife, seen antenatally, or in labour. Due to limitations, these process of care outcomes were not presented in this abstract
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Application of rigid exclusion criteria based on previous obstetric complications and previous medical history allowed low-risk cases only into study. Previous obstetric complications: Caesarean section; difficult forceps delivery; shoulder dystocia; anal sphincter tear; severe post-partum haemorrhage; pre-term delivery; intrauterine growth retardation; severe pre-eclampsia/eclampsia; perinatal loss and habitual abortion. Previous medical history of significant medical disorder: cardiovascular disease/ diabetes mellitus and gestational diabetes/ chronic renal disease/ autoimmune disease; drug addiction; alcohol abuse; long-standing infertility. 3 1000 women randomised: 495 to team care; 505 to standard care. After transfers, miscarriages, terminations, move to other hospital: 464 in team care (475 babies); 471 in standard care (466 babies) 4 For remainder of gestation starting at point of recruitment prior to first medical check-up until delivery (a minimum of 15 weeks follow-up in pregnancy) plus two months post-natal 5 Data on procedures in antenatal, intrapartum and postnatal periods plus data on maternal and infant care outcomes were extracted from medical records for mothers and infants. Maternal satisfaction was self-reported by follow-up questionnaire two months after discharge.

Results Quantitative results	*Satisfaction with antenatal care: team care vs. standard care – odds ratio (95% CI)* I was always kept informed about what was happening and doctors/midwives made an effort to explain anything I didn't understand: 2.17 (1.64–2.88) I was always given an active say in decisions about care in pregnancy: 1.49 (1.13–1.97) The doctors/midwives were encouraging and reassuring: 2.46 (1.84–3.29) Often at my check-ups the doctors/midwives were very rushed: 0.26 (0.20–0.34) Care in pregnancy was provided in a safe and competent way: 2.17 (1.63–2.89) I was happy with the physical aspects of care during pregnancy by doctors/midwives: 2.02 (1.53–2.68) I was happy with the emotional support I received in pregnancy by doctors/midwives: 2.39 (1.81–3.16) Overall, care during pregnancy was very good: 2.22 (1.66–2.95) *Satisfaction with intrapartum care: team care vs. standard care – odds ratio (95% CI)* I was always kept informed about what was happening and doctors/midwives made an effort to explain anything I didn't understand: 1.69 (1.28–2.23) I was always given an active say in decisions about care during labour and birth: 1.65 (1.25–2.17) The doctors/midwives were sensitive and understanding: 2.07 (1.56–2.76) The doctors/midwives were encouraging and reassuring: 1.85 (1.39–2.48) I often felt the doctors/midwives were very rushed: 0.61 (0.47–.81) Care during labour and birth was provided in a safe and competent way: 1.93 (1.43–2.59) I was happy with the physical aspects of care by doctors/midwives: 1.94 (1.46–2.59) I was happy with the emotional support I received from doctors/midwives: 1.78 (1.34–2.38) My needs of privacy were well respected during the birth: 1.91 (1.42–2.57) *Satisfaction with postnatal care: team care vs. standard care – odds ratio (95% CI)* I was always kept informed about what was happening and doctors/midwives made an effort to explain anything I didn't understand: 1.32 (1.01–1.73) I was always given an active say in decisions about care of my baby and myself: 1.20 (0.91–1.57) I was given the advice and support I needed in breastfeeding: 1.08 (0.82–1.42) I was given the advice and support I needed in how to handle, settle or look after the baby: 1.09 (0.84–1.43) I was given the advice and support I needed in any problems with the baby's health and progress: 1.10 (0.84–1.44) I was given the advice and support I needed about my own health and recovery after the birth: 1.16 (0.88–1.51) The midwives/doctors were sensitive and understanding: 1.34 (1.02–1.76) The midwives/doctors were encouraging and reassuring: 1.42 (1.09–1.87) I often felt the doctors/midwives were very rushed: 0.73 (0.56–1.05) Care in hospital after the birth was provided in a safe and competent way: 1.22 (0.93–1.60) I was happy with the physical aspects of care by doctors/midwives: 1.42 (1.08–1.86) Overall, the care in the hospital after birth was very good: 1.27 (0.97–1.67) When asked for the preference for care provider in antenatal period in the event of a new pregnancy, 50.3% in the team care group said they would prefer midwives only, compared with 21.8% in the standard care group, and 2.8% said they would prefer doctors only with 18.7% respectively. Fewer women in team care did not mind whether the care provider was a doctor or a midwife.

	Summary The team midwife care had the greatest impact in the antenatal period followed by the intrapartum period and then the postpartum period. In all measures in the antenatal period women in the team group were more satisfied with antenatal care. Women in the team group were generally more satisfied with care received during labour and delivery, though the differences in the groups were slightly less pronounced than for care during the antenatal period. Postnatal care did not differ greatly between the two groups, except for visits by a team midwife in the team care group that centred around the mother's feelings and well-being. There was no statistical difference in the overall assessment of postnatal care, but mothers in the team care group felt better informed, and perceived their care providers as more sensitive, understanding, encouraging, reassuring, and less rushed. They were also happier with the physical aspects of care provide by doctors and midwives. 84% of mothers in the team care group would choose the same model of care in a future pregnancy compared to 60% of the standard group.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No case mix adjustment reported. 2 No statistical difference in age; gestation at booking; parity; marital status; English as 1st language; education; family income; smoking 3 Gap in antenatal care data recognised for participants under shared care with community-based practitioners and for participants transferred away. 73% team care group responded to 2-month follow-up questionnaire; 64% standard care group responded to 2-month follow-up questionnaire. 4 Data analysed by intention to treat, ignoring moving away/transfers. *During recruitment* Team care: 4% premature end to pregnancy + 2.2% moved away/transferred; 6.6% lost to follow-up Standard care: 5% premature end to pregnancy + 1.8% moved away/transferred; 6.8% lost to follow-up *After randomisation and during follow-up* Team care: 2.2% moved to standard care + 4.5% moved away/transferred; 6.7% lost to follow-up Two sets of twins born adding four extra infant participants Standard care: 0% moved to team care + 5.5% moved away/transferred; 5.5% lost to follow-up Four sets of twins born adding two extra infant participants 5 Random allocation to team care or standard care using opaque numbered envelopes 6 One women's hospital in Victoria
Commentary	Linked to another study on the same population. A possible limitation is the dilution of the team midwifery intervention caused by inclusion of women receiving birth centre care within the standard care group – the two are similar as recognised by the authors. Sociodemography is well reported and the groups were similar. It may be that what is important to women are characteristics of individual encounters with caregivers, but having a known caregiver may be a means to an end. The team group may have differed from their colleagues because of their philosophy of care since they volunteered to take part in the trial. Continuity of care affected women's satisfaction with care positively and the authors believe that increased satisfaction with intrapartum and postpartum care was an effect of the continuity of midwife caregiver.
Research implications	A useful area of research is in the continuity of care. The elements that make women satisfied with the care they receive should be identified in future research with an emphasis on continuity.

ID, origin, authors (year)	168, Australia; Waldenström, U. *et al.* 2001
Aims	To add to the current literature on intervention rates and maternal and infant outcomes of a new model of team midwife care *Workforce:* Secondary care antenatal clinic setting Nursing workforce: midwife *Feature:* Substitution of midwife team care for standard models of care; operation of workforce with team work amongs midwives Intervention/comparison: A team of 8 midwives providing antenatal, intrapartum and postnatal care in collaboration with medical staff using the same medical protocols as standard care. Standard models of care included: (i) care from doctors mostly in a hospital clinic; (ii) care from midwives mostly in a midwives clinic (in collaboration with hospital doctors); (iii) birth centre care mostly run by midwives; (iv) shared care between general practitioners and hospital doctors. *Outcomes:* *Maternal* Maternal care outcomes: antenatal complications by type; gestation at delivery; duration of labour by stage; complications of labour and delivery by type; perineal status by type; length of postnatal stay Process of care outcomes: number of antenatal visits by type; number of ultrasound scans; number of antenatal CTG tests; number of visits to emergency care department; antenatal admissions by length of stay; met with midwives before labour (continuity of care measure); stage team midwife was present; status of accoucheur Procedures during labour: fetal monitoring by type; augmentation/induction/analgesia by type; operative procedures assisting birth by type Maternal problems reported two months into the postnatal period: soreness after tear or episiotomy; pain from Caesarean wound; incontinence; bowel problems; feeling tired and exhausted; more minor illnesses than usual; backache; sore nipples; mastitis, without medical treatment; mastitis treated with antibiotics; feeling depressed for more than a few days; constantly reliving labour; other *Infant* Infant care outcomes: perinatal mortality: number of stillbirths; number of neonatal deaths Measures of infant morbidity: admission to special care nursery (SCN); days in SCN; reasons for admission to SCN; number of pre-term babies; number of intrauterine growth retardation babies; Apgar score Due to the large number of reported outcomes, maternal and infant outcomes only will be fully presented as the health outcomes of interest.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Application of rigid exclusion criteria based on previous obstetric complications and previous medical history allowed low-risk cases only into study. Previous obstetric complications: Caesarean section; difficult forceps delivery; shoulder dystocia; anal sphincter tear; severe postpartum haemorrhage; pre-term delivery; intrauterine growth retardation; severe pre-eclampsia/eclampsia; perinatal loss and habitual abortion Previous medical history of significant medical disorder: cardiovascular disease; diabetes mellitus and gestational diabetes; chronic renal disease; autoimmune disease; drug addiction; alcohol abuse; long-standing infertility 3 1000 women randomised: 495 to team care; 505 to standard care. After transfers, miscarriages, terminations, move to other hospital: 464 in team care (475 babies); 471 in standard care (466 babies) 4 For remainder of gestation starting at point of recruitment prior to first medical check-up until delivery (a minimum of 15 weeks follow-up in pregnancy) plus two months postnatal 5 Data on procedures in antenatal, intrapartum and postnatal periods plus data on maternal and infant care outcomes were extracted from medical records for mothers and infants. Maternal problems in postnatal period were self-reported by questionnaire.

Results Quantitative results	**Maternal outcomes** *Maternal care outcomes: team care vs standard care – odds ratio (95% CI), unless otherwise stated* Antenatal complications: diabetes 0.0 (0.0–0.86); gestational diabetes 0.59 (0.55–1.37); antepartum haemorrhage 2.06 (0.77–5.69); pregnancy-induced hypertension 0.89 (0.29–2.71); mild pre-eclampsia 1.02 (0.51–2.03); moderate pre-eclampsia 0.51 (0.01–9.77); severe pre-eclampsia 0.58 (0.12–2.30); other– 0.63 (0.16–2.21) Gestation at delivery 2.3% vs. 2.0% (p = 0.66) Duration of labour: 1st stage (hours) 5.8 vs. 6.2 (p = 0.17); 2nd stage (minutes) 49.5 vs. 53.9; (p = 0.21); 3rd stage (minutes) 8.1 vs. 9.4 (p = 0.90) Complications of labour and delivery: shoulder dystocia 0.67 (0.14–2.86); prolapsed cord 0.0 (0.0–5.4); ruptured uterus 0% vs. 0.2%; postpartum haemorrhage (no Caesarean section) 1.08 (0.51–2.28); 3rd-degree tear 1.36 (0.23–9.31); post-Caesarean section bowel obstruction 0.2% vs. 0%; post-Caesarean section pulmonary oedema 0.02% vs. 0%; post-Caesarean section atelectasis 0.2% vs. 0% Perineal status: episiotomy 1.0 (0.74–1.35); sutured tear 0.67 (0.49–0.92); unsutured tear 1.27 (0.78–2.07); perineum intact 1.31 (0.96–1.8) Average length of postnatal stay (days) 3.8 vs. 3.7 *Relevant process of care outcomes: team care vs standard care – odds ratio (95% CI), unless otherwise stated* Fetal monitoring: auscultation 0.76 (0.53–1.08); CTG 0.81 (0.62–1.07); scalp pH 0.78 (0.36–1.68) Augmentation: 0.94 (0.69–1.26) Induction: 1.03 (0.78–1.37) Operative procedures: forceps 0.9 (0.62–1.32); vacuum extraction 0.75 (0.33–1.71); manual removal of placenta 0.6 (0.24–1.48) Caesarean section: elective 2.41 (0.86–7.72); emergency 0.82 (0.52–1.29) *Maternal problems reported two months into the postnatal period: team care vs standard care – odds ratio (95% CI), unless otherwise stated* Soreness after tear or episiotomy 0.81 (0.54–1.22); pain from Caesarean wound 0.66 (0.21–1.99); incontinence 0.9 (0.57–1.41); bowel problems 1.12 (0.65–1.94); feeling tired and exhausted 0.92 (0.67–1.27); more minor illnesses than usual 1.24 (0.74–2.08); backache 0.76 (0.55–1.05); sore nipples 0.76 (0.55–1.05); mastitis, without medical treatment 1.48 (0.78–2.84); mastitis treated with antibiotics 0.98 (0.57–1.67); feeling depressed for more than a few days 1.13 (0.77–1.65); constantly reliving labour 1.15 (1.67–1.96); other 1.89 (0.59–1.36) **Infant outcomes** *Infant care outcomes: perinatal mortality: team care vs standard care – n* Stillbirth: 4 vs. 4 Neonatal death: 1 vs. 3 *Infant care outcomes – infant morbidity: team care vs standard care – odds ratio (95% CI), unless otherwise stated* Admission to special care nursery (SCN) 1.4 (0.87–2.26) Mean days in SCN at >5 days 11.1 vs. 17.2 (p-value = 0.33) Reasons for admission to SCN: prematurity 1.02 (0.37–2.82); intrauterine growth retardation 1.53 (0.17–18.42); congenital malformation 0.2% vs. 1.1%; birth asphyxia 0 .4% vs. 0%; other 0.2% vs. 0.4% Total number of pre-term babies: 1.37 (0.82–3.11) Total number of intrauterine growth retardation babies: <10%, 1.59 (0.82–3.11); <3%, 1.17 (0.38–3.6) Apgar score: 1.32 (0.54–3.95)

	Summary Team midwife care did not reduce medical interventions in the study – interventions were fewer, but not significantly. Maternal health outcomes were very similar between the groups and no statistical difference were observed in infant outcomes. Overall, the trial showed no statistical differences between the team midwife care and standard care in medical interventions, maternal health and infant health. The figures on perinatal mortality suggest that team midwife care is not associated with a reduction in safety.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersa	1 No case mix adjustment reported. 2 No statistical difference in age; gestation at booking; parity; marital status; English as 1st language; education; family income; smoking 3 Gap in antenatal care data recognised for participants under shared care with community-based practitioners and for participants transferred away. 73% team care group responded to 2-month follow-up questionnaire; 64% standard care group responded to 2-month follow-up questionnaire. 4 Data analysed by intention to treat, ignoring moving away/transfers. *During recruitment* Team care: 4% premature end to pregnancy + 2.2% moved away/transferred; 6.6% lost to follow-up Standard care: 5% premature end to pregnancy + 1.8% moved away/transferred; 6.8% lost to follow-up *After randomisation and during follow-up* Team care: 2.2% moved to standard care + 4.5% moved away/transferred; 6.7% lost to follow-up Two sets of twins born adding four extra infant participants Standard care: 0% moved to team care + 5.5% moved away/transferred; 5.5% lost to follow-up Four sets of twins born adding two extra infant participants 5 Random allocation to team care or standard care using opaque numbered envelopes 6 One women's hospital in Victoria
Commentary	A possible limitation is the dilution of the team midwifery intervention caused by inclusion of women receiving birth centre care – the two are similar as recognised by the authors. There is also a gap in the antenatal data, due to poor access to some maternal medical records. No discussion of case mix adjustment for babies, although rigid exclusion criteria may negate this requirement in mothers. Inclusion of smoking status may help adjust for health status. Demography is well reported and the groups were similar. The hypothesis – that team midwife care does not reduce the incidence of medical interventions in this study – was not supported, although the authors make reference to other studies that do show this result.
Research implications	A meta-anlaysis of well-selected trials would help to determine whether team midwifery care is more beneficial than standard care. Continuity of care in pregnancy is important for expectant mothers and is achievable with a team care approach. The study could be repeated with a continuous care element and the result could help determine the critical factors that make team midwifery service a success.

ID, origin, authors (year)	270, Zimbabwe, Manungo, P.N. *et al.* (1996)
Aims	To evaluate the perinatal mortality and to describe how it is affected by nurse aide conducted deliveries, and to assess the impact of training *Workforce:* Secondary care setting: rural mission hospital Nursing workforce: nurse aides *Feature:* Substitution of nurse aides for doctors and nurses *Intervention/comparison:* The study examines referrals between and compares the outcomes of normal deliveries attended by nurse aides with no formal training to those primigravidae and complicated deliveries attended by trained staff of doctors and nurses. *Outcomes:* *Infant* Perinatal mortality rate
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Retrospective cohort study 2 No criteria stated 3 Total sample size of 1459 deliveries. Trained staff: n=4 (1 doctor + 3 registered nurses). Nurse aides: n=24 4 From delivery until infant discharge 5 Maternal data came from a review of hospital maternity records over period January 1992 until 1994. The hospital delivery book, T8 records and the hospital death register yielded data for number of deliveries, stillbirths, perinatal deaths and birth attendants
Results Quantitative results	**Infant outcomes** *Table of perinatal deaths by birth attendant* **Staff at delivery** — **Total deliveries (%)** — **Stillborn** — **Early neonatal deaths*** — **Perinatal mortality/1000†** Doctor/nurse — 635 (43) — 13 — 23 — 57 Nurse aide — 824 (57) — 1 — 3 — 5 Total — 1459 — 14 — 26 * first week of life † stillbirths + first-week deaths ÷ number of deliveries Pre-term 122; vacuum extractions 60; Caesarean sections 5; postpartum haemorrhage 9; stillbirths 14; early neonatal deaths: 26 **Summary** The referral between nurse aides and trained staff works well as supported by the low rate of perinatal mortality amongst nurse aide deliveries.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment reported. 2 Not reported. 3 No gaps in data collection reported. 4 Not applicable 5 Not applicable 6 One hospital at Mbuma, Nkayi District
Commentary	Examples of hospital maternity reports were provided. No clinical characteristics or demographics provided. No tests for significance presented. Reportage brief and incomplete. A brave effort by the author to present the findings of a study conducted into the outcomes of a substitution policy implemented as a last resort due to an acute shortage of trained staff.
Research implications	An important study for the region evaluating the effects of substitution in an acute nursing shortage. Further hospital reviews for other African district sites instituting similar policies should be conducted regionally for aiding health system planning.

Table A2.9 Skill mix

ID, origin, authors (year)	1209, USA, Aiken, L.H. *et al.* (2003)
Aims	To examine whether the proportion of hospital registered nurses (RNs) educated at the baccalaureate level or higher is associated with risk-adjusted mortality and failure to rescue (deaths in surgical patients with serious complications) *Workforce:* Hospital RNs *Feature:* Specialization of workforce *Outcome:* Risk-adjusted mortality and failure to rescue
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective, cross-sectional analyses 2 Include general, orthopaedic, and vascular surgery patients discharged from 168 (80% of 210) non-federal adult general Pennsylvania hospitals between 1 April 1998 and 30 Nov 1999. Exclude 26 hospitals because of missing data, most often because their reporting to external administrative sources was done as aggregate multi-hospital entitles. Exclude 6 Veterans Affairs hospitals because they do not report discharge data to the state. 10 small hospitals were excluded because most of them had 50 or fewer beds, and had an insufficient number of nurses responding to the questionnaire. 3 232,342 patients from 168 hospitals in Pennsylvania and 10,184 nurses 4 N/A 5 Patient discharge data were obtained from the Pennsylvania Health Care Cost Containment Council. Patient complications were determined with International Classification of Diseases, 9th edition. A 50% random sample of RNs residing in Pennsylvania received questionnaires at their homes and the response rate was 52%. Nurses were asked to indicate whether their highest credential in nursing was a hospital school diploma, an associate degree, a Bachelor's degree, a Master's degree or another degree. Nursing workload and number of years of experience working as an RN for nurses from each hospitals were also calculated.
Results Quantitative results	The proportion of hospital RNs holding a Bachelor's degree or higher ranged from 0% to 77% across the hospitals. After adjusting of patient characteristics and hospital structural characteristics (size, teaching status, level of technology), as well as for nurse staffing, nurse experience, and whether the patient's surgeon was board certified, a 10% increase in the proportion of nurses holding a Bachelor's degree was associated with a 5% decrease in both the likelihood of patients dying within 30 days of admission and the odds of failure to rescue (odds ratio 0.95; 95% Confidence Interval 0.91–0.99 in both cases).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Adjustments were made controlling patient characteristics and hospital structural characteristics (size, teaching status, level of technology), as well as for nurse staffing, nurse experience, and whether the patient's surgeon was board certified. 3 Yes 4 N/A 5 A 50% random sample of RNs received questionnaires; however, the response rate was 52%. Included 168 (80%) of adult general hospitals in Pennsylvania. 6 Pennsylvania
Commentary	This study provides the first empirical evidence that hospital's employment of nurses with BSN and higher degrees is associated with improved patient outcomes. One limitation of the analysis is the potential for response bias in the education and staffing measures derived from the nurse survey, with a 52% response rate. A second limitation related to study design. Longitudinal data sets, preferably including hospitals from more than one state, will be essential for establishing the generalisability of these findings.
Research implications	This study should be repeated using patient and nurse data from more than one state so results can be more generalizable.

ID, origin, authors (year)	529, USA, Anderson, R.A., Hsieh, P.C. and Su, H.F. (1998)
Aims	To identify patterns of resource allocation that relate to resident outcomes in nursing homes. *Workforce:* Nursing homes, registered nurses (RNs); tertiary care *Feature:* Skill mix and other characteristics in nursing homes: structure, human resource allocation, and financial resource allocation. According to the residents' health outcome, nursing homes were divided into two groups: group 1 – homes in the 80th percentile or higher having the best average resident outcomes; group 5 – home in the 20th percentile or lower having the worst average resident outcomes. *Outcome:* Verbal aggression, physical aggression, other disruptive behaviour, geriatric-chair restraints, wrist-mitten restraints, vest-belt restraints, contracture, decubitus ulcer, dehydration, urinary tract infraction, and fracture within the preceding 3 months.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional study 2 N/A 3 494 nursing houses 4 One year, in institution 5 Data on nursing homes were from the Texas department of Human Services (TDHS). Data on staffing levels and financial resource allocations were from the Texas Medicaid Nursing Facility 1990 Cost Reports (TDHS, 1990a). Data on case mix and resident outcomes were from 1990 Client Assessment, Review, and Evaluation Form 3652-A (TDHS, 1990b). Fiscal year 1990.
Results Quantitative results	Higher level and types of nursing staff relate to better quality of care. Raw means show that the group with the best average outcomes had a greater percentage of RNs, lower percentage of LVNs and nurse aides (NA) in the staff mix.

	Average Resident Outcomes: M(SD)		**Improvement in resident outcomes: M(SD)**	
	Best (n=97)	***Worst (n=96)***	***Best (n=97)***	***Worst (n=96)***
No. of RNs/60 beds	0.94 (1.33)	0.86 (0.82)	0.77 (0.77)	0.99 (0.92)
RN hours/resident day	0.10 (0.11)	0.10 (0.11)	0.08(0.08)	0.11 (0.10)
No. of RNs/total nursing staff	**0.03 (0.02)**	**0.02 (0.03)**	**0.02 (0.02)**	**0.03 (0.03)**
No. of LVNs/60 beds	6.03 (2.73)	6.38 (2.87)	6.06 (2.45)	5.69 (2.49)
LVN hours/resident day	0.60 (0.19)	0.63 (0.18)	0.62 (0.19)	0.59 (0.18)
No. of LVNs/total nursing staff	**0.25 (0.07)**	**0.26 (0.07)**	**0.26 (0.06)**	**0.24 (0.07)**
No. of NAs/60 beds	16.86 (6.17)	16.16 (6.05)	16.03 (5.09)	15.92 (5.91)
NA hours/resident day	1.68 (0.40)	1.60 (0.47)	1.64 (0.42)	1.68 (0.43)
No. of NAs/total nursing staff	**0.76 (0.08)**	**0.77 (0.07)**	**0.77 (0.07)**	**0.76 (0.07)**

Multivariate analysis of variance (MANOVA) was used to examine comparison groups for differences in staffing patterns for RNs, LVNs, and nurse aides. For the comparison between groups with Best and Worst Average Resident outcomes, all comparisons of mean scores were no significant difference; the same for the comparison between groups with Most and Least Improvement in Resident outcomes, but univariate F tests showed that RN pattern scores differed significantly, F (1, 192) = 5.08, p = 0.03, effect size (eta-squared) = 0.026, whereas LVN pattern scores and nurse aide pattern scores did not differ.

The residual scores (after case mix variance removed) for the 11 indicators of percentage of improvement in resident outcomes for groups with highest and lowest RN staffing levels showed that homes in the group with the highest levels of RN staffing has a generally higher improvement that the group with the lowest levels of RN staffing, except for the verbal aggression, vest and wrist restraints, contractures, and dehydration. Results of the ANOVA model showed that percentage of improvement in resident outcomes was significantly greater in homes with the higher levels of RN staffing, F (1, 191) = 7.06, p = 0.0009, effect size (eta-squared) = 0.038.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for the patients' age, gender, and 11 health indicators: mobility/ambulation problems, dressing/grooming problems, transferring problems, eating problems, toileting problems, bowel control problems, bladder control problems, changes in level of consciousness, dyspnoea, oedema, stasis ulcer, internal bleeding, and terminal illness 2 Adjusted for nursing home structure: owner status, number of licensed beds, and percentage of private pay 3 Uniformly reported 4 Complete 5 N/A 6 One state
Commentary	The study did not include psychosocial indicators of resident outcomes and, thus, did not fully capture the quality of care in these nursing homes. Data were from one state, limiting generalisability. The secondary data has a substantial lag time between actual collection of the data and when data become available for secondary use. Large databases used in this study require substantial processing before they can be used to answer research questions. The study has carefully controlled for the influence of case mix on outcomes. Applied the configurational approach, pattern scores were developed that synthesised multiple indicators of resident outcomes to be analysed as a whole. Regrouped the sample in different ways to confirm the veracity of the initial findings. Used outlier analysis by choosing cases based on extreme scores, the range of scores used in analysis is restricted, and extreme values in the sample may represent sampling error, measurement error, or a misspecified model.
Research implications	Comparative studies of other states will add to knowledge about how resource allocation influences resident outcomes. Development of the RNs should make them more valuable; continuing education to improve RNs' skills is a logical investment. It might be beneficial to structure clinical experiences to facilitate students in learning how to gather and analyse group-level clinical data and to plan group-level interventions. More strict and comprehensive case mix adjustment is recommended to future studies. It is suggested that regrouping the sample in different ways will provide stronger evidence for making valid conclusions.

ID, origin, authors (year)	17, USA, Barkell, N.P., Killinger, K.A. and Schultz, S.D. (2002)
Aims	To explore the effects of a change in nurse staffing model on outcomes in postoperative bowel procedure patients. *Workforce:* Registered nurses (RNs) and patient care associates (PCAs); secondary care *Feature:* Skill mix: RN staffing model. Model A: a team nursing model with PCA assisting RN in delivery of patient care. The role of the RN was to direct and oversee patient care, delegating basic patient care activities such as bathing, feeding, ambulating, and turning patients. Model B: The RN is responsible for giving total care to the patient, the PCA's role included discharging and transporting patients, answering call lights, gathering equipment, and assisting the RN with activities that required two persons such as the ambulation or transfer of some patients. It halved the numbers of PCAs, and decreased 2 RNs. *Outcome:* Variable cost, length of stay (LOS), the incidence of pneumonia, the incidence of urinary tract infection (UTI), patient satisfaction, the patient's perception of pain as measured by the mean pain scores for postoperative days 1 and 2, and the frequency of documentation of pain scores for postoperative days 1 and 2.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Before-and-after study 2 Included patients between the ages of 18 and 85, who had bowel procedures with no pneumonia or UTI preoperatively, and the entire LOS having occurred on the study unit. 3 59 patients for staffing model A and 37 patients for staffing model B; patient satisfaction sample sizes were 139 and 108 for staffing models A and B respectively; patients are in a 33-bed inpatient surgical unit of a community-based teaching hospital. 4 1 year, in-hospital 5 Data were from medical records; clinical pathways and the nurses' narrative notes were used to retrieve patient pain scores and the number of pain scores documented by the nurse for postoperative days 1 and 2; patient satisfaction data were from the Parkside Patient Satisfaction Survey; variable cost and LOS were from the facility's information system. From June 1999 (model A) to June 2000 (model B); the time periods were chosen to avoid known variant factors that may have interfered with the results, such as employee peak vacation time, unit closure during December holidays, and the July influx of graduate nurses and resident physicians.
Results Quantitative results	There were few significant differences in patient outcomes between staffing model A and staffing model B. Using *t*-test, the differences between staffing model A and staffing model B with respect to variable cost, LOS, and patient satisfaction were not found to be statistically significant. The mean pain scores are significantly higher in model B, and the mean number of pain scores that were significant were lower in model B. No UTIs occurred in either group, and pneumonia occurred in only 5.1% of patients in model A and in none of the patients in model B, thus no other statistical analysis was done. *Comparison of dependent variables in staffing model A and staffing model B: Mean (SD)*

	Staffing model A	**Staffing model B**	***p*-value**
Length of stay	6.8 days (3.1)	7.1 days (2.9)	0.627
Patient satisfaction score*	83.4 (12.8)	84.6 (12.8)	0.468
Number of documented pain score **	7.5 (2.67)	6.0 (2.79)	0.006
Pain score**†	1.9 (0.9)	2.6 (1.32)	0.017

* on a scale of 0 to 100
** postoperation days 1 and 2
† on a scale of 0 to 10
$p < 0.05$

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 The groups were found to be equivalent on age, demographic characteristics, comorbidities, and primary diagnoses; therefore no further adjustments were considered. 2 N/A 3 Uniform 4 Complete 5 Convenience sampling 6 One hospital, single setting
Commentary	The study took place within the context of a rapidly changing health care environment. Intervening variables, including bed closures and re-openings without staff re-adjustments and budget reductions, may have affected the results. However, it occurred in a 'real world' setting with limited ability to exercise control over multiple factors that influence outcomes.
Research implications	Since the nursing staff's ability to achieve comparable outcomes in light of decreased numbers of caregivers may have been due in part to their investment in a change that they had initiated, examination of job satisfaction, absenteeism, and staff turnover may have yielded important insight into the effects of staffing model changes on nursing staff. More researches are needed that control for patient acuity and actual time spent in direct care. The use of larger and more diverse patient populations will facilitate generalisability of the findings. More studies are needed to address the relationship between nurse staffing and pain management. Qualitative studies may unveil worthwhile information regarding patient satisfaction with nursing care. This study can help nurse administrators determine the appropriate number and skill level of nursing staff needed to provide safe, high-quality patient care. Staffing level must correspond to the needs of the patients on each unit and facilitate the achievement of desired outcomes.

ID, origin, authors (year)	194, USA, Bolton, L.B., Jones, D., Aydin, C.E. *et al.* (2001)
Aims	To explore the need for evidence-based health policy, as illustrated by the mandatory staffing bill passed by the California state legislature in 1999 *Workforce:* Registered nurses, licensed vocational nurses and other caregivers; secondary care *Feature:* Nursing staff mix and hours of care (productive hours worked by the nursing staff who provide direct patient care on the defined unit and are included in the staffing matrix) *Outcome:* Falls and Pressure ulcers
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional, observation 2 Those hospitals who joined the California Nursing Outcomes Coalition (CalNOC) 3 257 units (30 medical, 29 surgical, 73 medical–surgical combined, 55 step down, 65 critical care and 5 24-hour observation units) from 38 hospitals, representing 1,253,892 inpatient days 4 In-hospital 5 Data were collected about patients or units from June 1998 to June 1999. Staff reported direct hours of care using hospital information systems and patients' falls using 'incidence' databases. Data were then extracted by hospital personnel and entered into the (CalNOC) database. Clinical staff assessed the prevalence of pressure ulcers and recorded patients' observations for submission to CalNOC.
Results Quantitative results	Hospital (pressure ulcer analyses) and unit (falls analyses) levels indicated that hospitals and units where patients received >70% of their care from RNs had similar rates of falls and hospital-acquired pressure ulcers as did hospitals where less than 50% of care was provided by RNs. Early data (28 hospitals) did not indicate an association between the hours of care provided by RNs or skill mix of patient care providers and the occurrence of patient falls or prevalence of pressure ulcers.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 None reported. 2 None reported. 3 Yes 4 Complete 5 Convenience sampling 6 9% of all general acute hospitals in California
Commentary	No adjustment for case mix or other potential confounders.
Research implications	There is a need for extensive analysis to determine safe minimal staffing ratios. Research is needed for examining evidence of the association between errors and numbers of RNs and other professionals. More research is needed on the variables contributing to error reduction. Data are needed that: (i) provide clear definitions of the level of expected quality; (ii) relate patient outcomes to staffing at the unit-type level; (iii) use appropriate measures of risk adjustment.

ID, origin, authors (year)	476, USA. Brett, J.L.L., and Tonges, M.C. (1990)
Aims	To test the effect of the Professional Advanced Care Team model (ProACT), aimed at enabling few RNs to provide quality nursing care, on the quality of care, costs of care, and patient and staff satisfaction. *Workforce:* Registered nurses (RNs); secondary care *Feature:* Nursing skill mix, nursing workload, and role expanding among RNs, licensed practice nurses (LPNs), and nurse aides (NAs) *Outcome:* Quality of care: compliance; the percentage achievement of nursing process criteria; the percentage achievement of nursing outcome criteria; the number of incidents attributable to nursing care; the number of infections attributable to nursing care. Patient satisfaction; staff satisfaction. Costs of care.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Before-and-after study 2 Only piloted the model on a unique unit, no comparable control unit was available. 3 One 32-bed surgical orthopaedic unit; 25 patients during each evaluation period. 4 8 months; in-hospital 5 Patients' incident data were from routine incident reports prepared by the nursing staff for the Risk Management Department; infection data were from the Infection Control Department. Acuity statistics were from the Workload Management System for Nursing (WMSN). Patient satisfaction data were from a structured interview by Support Services Department, Discharged Patient Survey developed by the marketing department, and the Patient Satisfaction with Nursing Care Instrument. Pre-implementation evaluation period (T1): April and May 1988; post-implementation evaluation period (T2): September and October 1988, the fourth and fifth months of the model's existence. The third follow-up evaluation period (T3): January 1989, the eighth month of the model's existence.
Results Quantitative results	Fewer RNs with increased clinical and nonclinical support can provide high-quality, efficient care. Comparison of the mean score of patient satisfaction, number of incidents, infections per patient day, and compliance; a one-way analysis of variance of the patient satisfaction scores indicated that there was no significant difference (df, 32; p = 0.49). Thus despite the provision of care with markedly fewer RNs, patient satisfaction remained stable.

	T-1a	**T-2b**	**T-3c**
Patient satisfaction	3.1	3.4	3.4
Compliance – process criteria	63.3%	85.3%	77.6%
Compliance – process criteria	81.3%	89.4%	90.0%
Incidents/patient day	0.007	0.007	0.007
Infections/patient day	0.003	0.002	0.003

Quality appraisal	
1 Case mix adjustment	1 N/A
2 Other adjustment	2 N/A
3 Uniform data collection	3 Uniform
4 Participant follow-up	4 Complete
5 Random sampling	5 Convenience
6 Geographical dispersal	6 One unit, single setting
Commentary	No detailed information on included patients. The model was in place for less than one year and has been tested on only one nursing unit. The evaluation research focuses on outcomes rather than understanding why something works; the possibility exists that variables other than ProACT may be responsible for the results, i.e. without a control unit, the effect of the change process itself on the results is unclear.
Research implications	Need continued intermittent monitoring to establish the stability of the observed results. Implementation and evaluation on additional nursing units will be helpful. Restructuring hospital practice and care delivery to free nurses from non-nursing tasks is a useful approach to balancing the demand for nurses with the available supply.

ID, origin, authors (year)	734, UK, Carr-Hill, R.A. *et al.* (1995)
Aims	To analyse the relation between skill mix of a group of nurses and the quality of care provided (by Quality of Patient Care Scale QUALPACS: psychosocial care, physical care, general care, communication on behalf of the patient and professional implications) and outcomes (8 dimensions: patient hygiene, nutrition and hydration, skin integrity, intravenous therapy, planning for discharge, pain control, education/rehabilitation and elimination) *Workforce:* Nurses; Secondary *Feature:* Skill-mix *Outcome:* Patient hygiene, nutrition and hydration, skin integrity, intravenous therapy, planning for discharge, pain control, education/rehabilitation and elimination
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective, correlation study 2 Included nurses and patients in 15 wards in 7 hospitals who completed and returned forms and questionnaires. 3 Data were collected on 15 wards at 7 sites: 359 QUALPACS forms, 720 outcome measurement forms, 90 dependency and workload measure, 360 activity sample forms, 248 staff questionnaires 4 N/A 5 A trained observer watched all the interactions between nursing staff and two selected patients over 2-hour periods and used the 60-item QUALPACS check list to note the grade of staff and the quality of care delivered of each interaction. 24 of these 2-hour sessions were carried out on each ward during the 6 days the research team were in residence. At the end of the QUALPACS session, the nurse observers used two copies of the outcomes instruments (one for each patient) to make a summative assessment of the outcomes of the nursing care received in that session.
Results Quantitative results	The relationship between skill mix of a group of nurses and the quality of care provided and outcomes was examined. The correlation between the proportion of nurse staff on Grade D or above and QUALPACS scores (0.53, p = 0.02) is stronger than that between those staff and the achievement of good outcomes (0.30, p = 0.14).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No 2 The traditional approach to a multivariate analysis would be to include dummy variables for the hospital ward in a regression framework or in an analysis of variance. However, when the data are constructed in this hierarchial fashion, the usual assumptions of multivariate analysis is not tenable. Instead, variance is partitioned between the different levels. The standard error term is partitioned between an error attributable to the lowest level of analysis and an error attributable to the natural grouping of these units. Their variances can be separately estimated. 3 Observers who completed QUALPACS score and outcome forms may rate differently. Tests during the training of raters showed statistically significant differences between raters in the number of ratings they gave, but they did not show any statistically significant difference in average overall quality when measured over a QUALPACS session. But still this causes some unreliability on the scoring system. 4 N/A 5 Not random. The choice of settings was opportunistic and constrained by time and resources. 6 Not stated. Involve 7 hospitals.
Commentary	There was no separate instrument to collect data on nursing skill mix, so this study was limited to using grade as a proxy for skill. The choice of settings was opportunistic and constrained by time and resources so that the skill mix combinations studied were those that happened to be on the wards.
Research implications	Quality of nursing care improved as the ratio of qualified and further trained staff increased while costs increased with the quality of nursing care. Thus, employing qualified staff, providing post-qualification training and developing effective methods of organising nursing care need to be regarded as investments which pay dividends in the delivery of good-quality nursing care.

<table>
<tr><td>ID, origin, authors (year)</td><td>557, USA, Huston, C.J. (2001)</td></tr>
<tr><td>Aims</td><td>To examine correlations between staffing mix and pain management as a process indicator of quality after the implementation of a staffing model designed to increase unlicensed assistive personnel and decrease registered nurses and licensed vocational nurses in the skill mix.
Workforce: Registered nurses (RNs), unlicensed assistive personnel (UAP), and licensed vocational nurses (LVNs); secondary care
Feature: Skill mix: nursing staff models. Traditional team-leading model (first quarter of 1996): composed of RNs, , LVNs and UAP. New staffing model (first quarter of 1997): expand the use of UAP and reduce the number of RNs.
Outcome: Post-surgery pain: measured by numeric pain scale scores; 0 means no pain, 10 means intolerable pain.</td></tr>
<tr><td>Methods
1 Design
2 In-/exclusion
3 Sample size
4 Follow-up time
5 Data collection: source and period</td><td>1 Before-and-after study
2 Patients admitted to the two selected hospitals with a discharge diagnosis-related group of major joint and limb reattachment procedures of the lower extremity.
3 203 patients in 2 surgical units in 2 hospitals; 95 patients were admitted during the first quarter of 1996, 108 patients were admitted during the first quarter of 1997.
4 One year, in-hospital
5 Staffing mix data were from Productivity Reports issued by both hospitals; pain scale data, patients' demographics, and the type of surgical procedure performed were from a random chart review completed by patients; the patients' information was collected at 5 time points: exit from the postanaesthesia care unit (PACU), on admission to the postoperative unit, and the first documented pain scale scores during the first, second and third full shifts after surgery.</td></tr>
<tr><td>Results
Quantitative results</td><td>A positive correlation exists between RN staffing levels as a percentage of staffing mix and lower pain scores as reported by patient after surgery; while the opposite situation exists for UAP.
Comparison of mean pain scale scores between the first quarters of 1996 and 1997, according to different pain management styles (NAA, PCA, EA/SA and EA/SA/PCA*)
<table>
<tr><th>Pain score measurement point</th><th>Mean pain scale score (first quarter 1996)</th><th>Mean pain scale score (first quarter 1997)</th><th>t-value</th><th>p-value</th></tr>
<tr><td>Hospital A</td><td></td><td></td><td></td><td></td></tr>
<tr><td>PACU exit score</td><td>1.375 (n=4)</td><td>None recorded</td><td></td><td></td></tr>
<tr><td>Unit admit score</td><td>4.50 (n=10)</td><td>7.00 (n=5)</td><td>–1.346 (df = 13)</td><td>0.20</td></tr>
<tr><td>First shift</td><td>3.44 (n=17)</td><td>5.85 (n=10)</td><td>–2.501 (df = 25)</td><td>0.02</td></tr>
<tr><td>Second shift</td><td>3.27 (n=13)</td><td>5.50 (n=7)</td><td>–2.133 (df = 18)</td><td>0.05</td></tr>
<tr><td>Third shift</td><td>4.14 (n=14)</td><td>5.70 (n=10)</td><td>–1.664 (df = 22)</td><td>0.11</td></tr>
<tr><td>Hospital B</td><td></td><td></td><td></td><td></td></tr>
<tr><td>PACU exit score</td><td>3.67 (n=6)</td><td>1.95 (n=11)</td><td>1.584 (df = 15)</td><td>0.13</td></tr>
<tr><td>Unit admit score</td><td>4.25 (n=8)</td><td>5.20 (n=5)</td><td>–0.588 (df = 11)</td><td>0.57</td></tr>
<tr><td>First shift</td><td>4.44 (n=8)</td><td>6.50 (n=8)</td><td>–1.949 (df = 14)</td><td>0.07</td></tr>
<tr><td>Second shift</td><td>4.70 (n=5)</td><td>5.50 (n=5)</td><td>–0.645 (df = 8)</td><td>0.54</td></tr>
<tr><td>Third shift</td><td>4.64 (n=7)</td><td>4.90 (n=5)</td><td>–0.192 (df = 10)</td><td>0.85</td></tr>
</table>
* NAA: nurse-administered analgesia; PCA: intravenous patient-controlled analgesia; EA/SA: epidural/spinal analgesia; EA/SA/PCA: the combination of PCA and EA/SA</td></tr>
</table>

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No significant difference between the ages of patients during two study periods; there is no adjustment for patients' characteristics. 2 N/A; Subgroup analysis according to pain management types 3 Uniform 4 Complete 5 Random sampling of patients admitted to two hospitals 6 Two hospitals
Commentary	The use of a retrospective research design is subject to substantial limitations, i.e. the inconsistency of documentation by nursing staff. No adjustment on the patients' demographics.
Research implications	Future study is recommended to limit the sample population to a single diagnosis (DRG) to reduce extraneous variables. The breakdown of pain scale scores by primary type of pain management is strongly recommended for future research. Because nurse-administered analgesia (NAA) pain scores are more sensitive to direct nursing interventions and thus more sensitive to staffing mix changes than other types of pain management strategies, future study on defining the nursing sensitivity of pain management as an outcome indicator can limit sample to patients with NAA. Well-controlled concurrent or prospective studies may reduce the likelihood of the inconsistent or inadequate documentation. The identification and measurement of nursing-sensitive patient outcomes is little known, and quality of nursing care has yet to be defined. The specific nursing interventions that make a difference in patient outcomes have not been clearly identified; other factors contributing to the outcomes need to be studied. There is a need for continual reassessment of the validity of the structure, process, and outcome indicators currently recommended for use. Pain management is affected by a number of different inputs, and although nursing interventions are clearly a chief input, the weighted influence of each input is not fully understood. An action plan for the future should have three objectives: (i) maximum use of each health care member to deliver safe, effective, and appropriate patient care; (ii) increased productivity of both professional and ancillary staff; (iii) a contemporary and stable framework for nursing practice.

ID, origin, authors (year)	420, USA, Krainovich-Miller, B. *et al.* (1997)
Aims	To determine whether there has been any significant improvement in the quantity and quality of the science supporting increased delegation of nursing tasks to unlicensed assistant personnel (UAP) and if prior recommendations had been implemented. *Workforce:* Nurses; Secondary *Feature:* Skill-mix *Outcome:* Any patient outcomes
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Non-systematic review 2 Reported reviews of UAP research conducted between 1988 and 1994; the most recent UAP nursing research conducted between 1994 and 1997 3 N/A 4 Reviews (2); descriptive questionnaire (2); retrospective comparative (1); experimental pre-test/post-test (2) 5 CINHAL computer search 6 The studies included in the two reviews were frequently based on anecdotal 'evidence' and fraught with methodological limitations such as mismatched groups, small convenience sample sizes, and tolls of questionable validity and reliability. Among the 1994–1997 studies, two of them used convenience samples, two used single settings, two examined quality outcome indicators. 7 N/A
Results Quantitative results	There is very little research to substantiate institutional claims that these new systems of care can maintain quality and cut costs.
Commentary	Nurses should have an ethical and moral obligation to involve the researches on RN/UAP delivery care models. Multi-site and longitudinal studies are needed. Models of nursing care delivery studies, quality indicator studies are needed. The curricula of undergraduate and graduate programs, continuing education courses, and staff development training programs should be revised to include a Model of UAP Management, which include theories related to change, conflict resolution, leadership, and management as well as specific principles of delegation, supervision, performance review, and competency measurement.
Research implications	As a nature of evidenced-based literature review, not so many details of the studies have been reported; however, it focused on the research, education, and practice implications of all these studies.

ID, origin, authors (year)	1040, Northern Ireland, McKenna, H.P. (1994)
Aims	To explore the relationship between skill substitution and quality of care *Workforce:* Nurses *Feature:* staffing skill mix: a skill mix of mostly qualified staff ; a skill mix of mostly unqualified staff *Outcomes:* Patient satisfaction, mortality, length of stay (LOS), cost, staffs' moral, staffs' productivity and effectiveness
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	 1 Literature review 2 Inclusion: any studies that explore the relationship between skill substitution and quality of care. 3 – 4 – 5 – 6 – 7 –
Results Quantitative results	None
Results	11 studies suggested that 'a skill mix of mostly qualified staff is often an inefficient and ineffective way to run a health service'. 16 studies suggested that 'a skill mix of mostly unqualified staff is often an inefficient and ineffective way to run a health service'. 38 studies suggested that 'a skill mix of mostly qualified staff is a highly efficient and highly effective way to run a health service'.
Commentary	It is an evidence-based literature review; nothing has been mentioned about the quality of the relevant studies.
Research implications	Need high-quality replicate research in this area and practitioners must get involved in skill mix reviews and prove that they are efficient and effective in the myriad of new roles they will take on in the new health service.

ID, origin, authors (year)	80, USA, Needleman, J. *et al.* (2002)
Aims	To examine the relation between the amount of care provided by nurses at the hospital and patients' outcome *Workforce:* Nurses; secondary care *Feature:* Level of staffing by registered nurses, licensed practical nurses, and nurses' aides were estimated: the hours of nursing care per inpatient-day; the proportion of hours of nursing care provided by each category of nursing personnel. *Outcome:* 14 outcomes potentially sensitive to nursing (OPSN): length of stay (LOS); urinary tract infection (UTI); pressure ulcers; hospital-acquired pneumonia; shock or cardiac arrest (CA); upper gastrointestinal bleeding; hospital-acquired sepsis; deep venous thrombosis (DVT); central nervous system complications; in-hospital death; failure to rescue (FTR); wound infection; pulmonary failure (PF); metabolic derangement (MD). The last three adverse outcomes are only for surgical patients. Executive summary may be found at http://www.bhpr.hrsa.gov/nursing/
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional study 2 It excluded hospitals with an average daily census of less than 20, an occupancy rate below 20%, or missing data on staffing, as well as those reporting extremely low or high levels of staffing per patient-day. 3 5,075,969 medical patients and 1,104,659 surgical patients from 799 hospitals in 11 states 4 1 year; in-hospital 5 Hospital patient discharge data for OPSNs, and state hospital financial data or hospital staffing surveys for measures of nurse staffing at different levels; 1997–1998
Results Quantitative results	An inverse association between registered nurses and adverse outcomes, but not for licensed practical nurses or aides. *Medical patients:* Proportion RN-hours (*p*-value)/no. hours (*p*-value). LOS 1.12 (0.01)/–0.09 (<0.001), consistent*; UTI 0.48 (0.001)/0.99 (0.003), consistent; pneumonia 0.59 (0.001)/0.99 (0.08), consistent; shock 0.46 (0.007)/0.98 (0.22); bleed 0.66 (0.03)/0.98 (<0.007), consistent. No association or inconsistent relationship: ** Ulcer, inconsistent; sepsis, none; DVT, none/ inconsistent; CNS complications, none; death, none; FTR, inconsistent. *Surgical patients:* UTI 0.67 (0.04)/1.00 (1.00); pneumonia, weak; FTR 0.73 (0.12)/0.98 (0.008) consistent. No association or inconsistent relationship: ** LOS, none; ulcer, none; shock/CA, none; bleed, none; sepsis, none; DVT, none; CNS complications, inconsistent; death, none; wound, none; PF, none; MD, none. * Consistent: the changes are in the same direction among all the models of nursing ** Relationship is presented in Nurse Staffing and Patient Outcomes in Hospitals, Executive Summary by Needleman *et al.*, US Department of Health and Human Services Health Resources and Services Administration Contract No. 230-99-0021 (accessed 25 April 2003, at ftp://ftp.hrsa.gov/bhpr/nursing/staffstudy/staffexecsum.pdf)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Estimated the level of nursing care needed by patients in each diagnosis-related group to construct a nursing case-mix index for adjusted number of needed nursing hours per day; a patient-specific risk index based on patient diagnosis (DRG), the state of residence, age, sex, primary health insurer, whether or not the patient was admitted on an emergency basis, and the presence or absence of 13 chronic diseases. 2 Hospital characteristics including location, number of beds, occupancy rate, and teaching status were included in the analysis. Additional adjustments were made for patient acuity in each hospital's mix of patients. 3 The staffing levels of nurses for inpatient care from the diverse data sets of multiple states required substantial efforts to standardise the data and to determine what proportion of a hospital's nursing staff was allocated to inpatient care. 4 Complete follow-up 5 Unclear 6 26% representative

Commentary	The adverse outcomes are likely to be under-reported because of the inappropriate coding and exclusion rules for each adverse outcome. The smaller size of the samples of surgical patients may make it difficult to detect the association between staffing level and health outcomes. OPSNs were more than likely under-reported, did not include all possible outcomes of interest, and were biased toward adverse outcomes.
Research implications	Need to refine the measurement of the nursing care mix on the basis of discharge data and to elucidate the factors influencing the staffing levels of nurses and the mix of nursing personnel in hospitals. Systems should be developed for the routine monitoring of hospital outcomes that are sensitive to levels of staffing by nurses and ensure that an adequate nursing staff is available to protect patients and to improve the quality of care. Expand and refine OPSNs, including developing and testing measures of positive patient outmodes related to nursing. Update and refine the measurement of patient nursing acuity in discharge data sets. Understand the factors influencing both nurse staffing levels and mix of nursing personnel in hospitals, how these are influenced by case mix, bed mix, physical layout of hospitals, nursing practice patterns, market and financial pressures, and the availability of nursing personnel

ID, origin, authors (year)	1157, Australia, Pearson, A. *et al.* (1991)
Aims	To investigate the relationships between skill-mix, resident dependency and the quality of care and life in nursing homes *Workforce:* Nurses; tertiary care *Feature:* Skill mix: the qualifications of the staff (staffing mix: the proportion of qualified and unqualified staff); the exposure of staff to in-service training and the leadership style of the director of nursing *Outcome:* Quality of health care, resident's rights, social environment, and physical environment
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Excluded nursing homes with 20 residents or fewer 3 1374 patients in 200 non-governmental nursing homes in 4 states 4 14 months, in nursing homes 5 Data on staffing mix were from 200 staffing profiles; data on other skill mix measurements were from 3700 self-administered questionnaires for all staff and 199 director of nursing/charge nurse questionnaires; data on relative dependency (the relative residents' needs of service) were from 3345 Resident Classification Instruments; data on health outcome were from 200 observation schedules and 1374 resident interviews. Two stages: stage 1 from August 1988 to February 1989 – a pilot study aimed to develop instruments and procedures which would elicit the relevant data required; stage 2 from March 1989 to May 1990 – an analysis process
Results Quantitative results	The staffing mix, as measured by the type and level of training of nursing home staff, influences the level of care/life to a limited degree. The important aspects of skill mix are the commitment within the nursing home to in-service training and the leadership, management style and professional activity of the director of nursing. *Regressions of outcome measures on skill mix and dependency and home ownership and size* **Health care** / **Privacy and dignity** / **Freedom of choice** / **Variety of experience** / **Social independence** / **Home-like environment** Percentage RN: 0 / –0.001 / 0.001 / -0.003 / 0.005 / 0.002 Percentage EN: 0.002 / 0.002 / 0.001 / –0.001 / –0.003* / –0.003* Percentage therapists: 0.002 / -0.005 / –0.003 / 0.026* / –0.001 / 0.013 * Significant at the 5% level
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Adjusted for patients' dependency 2 Adjusted for the organisational variables: the size, type and location of home 3) N/A 4 Complete 5 Stratified random sampling; nursing homes are stratified by organisational variables. 6 4 states
Commentary	No adjustment of other patient characteristics than relative dependency No information on the uniform data collection The data sources for organizational variables were not reported.
Research implications	Further studies should pay attention to the increase in the proportion of therapists and the role and training of State Enrolled Nurses in the industry. It is difficult to produce defensible prescriptive percentages of optimal staffing mix; however, there are several stages suggested to move toward the development of optimal skill mix: maintain registered nurses at current levels; increase therapist, particularly in the area of diversified therapy, to at least the notional 8%; clarify the role of the enrolled nurse in nursing homes and review the role of nursing assistants.

ID, origin, authors (year)	309, Australia, Pratt, R. *et al.* (1993)
Aims	To compare the quality and costs of nursing care as assessed by patient outcomes of two different staffing regimes in each of two wards in one hospital *Workforce:* Registered nurse (RN) and enrolled nurse (EN); secondary care *Feature:* Skill mix – percentage of RNs: all-RN staffing regime; 80% RN, 20% EN staffing regime *Outcome:* Cost and patient outcomes – judgements by patients as to the quality of the nursing care they had received; estimations by qualified observers as to the completeness and accuracy of patients' nursing notes and records; observations by assessors as to the correctness with which a range of nursing procedures was performed and recorded.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 N/A 3 1 acute medical ward (17 beds) and 1 acute surgical ward (25 beds) in 1 hospital 4 3 months, in-hospital 5 For patients' dependency level, the Patient Assessment Information System (PAIS); for patient outcomes, self-designed questionnaires. First 8 weeks – all RN; next 4 weeks – RN and EN mix; final 8 weeks – RN and EN mix. (No specific year or months for study mentioned)
Results Quantitative results	There were comparatively few differences in patient outcomes between the staffing regimes on either ward. Patients were asked a wide range of questions related to nurses' responses to patients' needs in areas which included cardio-respiratory functioning, hydration and nutrition, elimination, comfort, communication and self-esteem, and safety and privacy. The results from the acute medical ward: 85% (50) patients did not have preference on the types of staffing, 12% (7) favoured the all-RN staffing, and 3% (2) favoured the RN-EN staff mix. In the acute surgical ward: 94%(83) patients did not have preference on the types of staffing, 6% (5) favoured the all-RN staffing, and none favoured the RN-EN staff mix.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No significant differences between patient dependency and staff productivity, bed occupancy rates and patients' duration on either ward, no further adjustment needs to be made. 2 N/A 3 Uniform 4 Complete 5 N/A 6 One hospital
Commentary	Only in one hospital. No details on the year or month of the study. No adjustment for the other characteristics of the wards, such as equipment and facilities of the wards.
Research implications	Longer-term studies need to be carried out by single or combined health care agencies. When ENs are introduced on to ward in the long term, then very careful monitoring of RN's burn-out, increased sick leave, and resignation rates should be built into evaluative criteria.

ID, origin, authors (year)	1130, Canada, Tourangeau, A.E., Giovannetti, P., Tu, J.V. and Wood, M. (2002)
Aims	To investigate the effects of nursing related hospital variables on risk-adjusted 30-day post-admission mortality rates for hospitalized patients. *Workforce:* Registered nurses (RNs); secondary care *Feature:* RN skill mix: RN inpatient earned hours proportionate to other inpatient nursing staff earned hours (RN, registered practical nurse, and unlicensed assistive personnel earned hours) Nursing dose: total inpatient clinical nursing worked hours per Ontario case weight (OCW) *Outcome:* 30-day mortality rates
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Included patients who were at least 20 years of age and had a diagnosis of acute myocardial infarction, stroke, pneumonia, or septicaemia as the initial reason for hospitalization. Excluded patients transferred from other acute-care hospitals, with a pre-admission or secondary diagnosis of cancer, palliative care, or immune deficiency disease. 3 46,941 patients discharged from 75 hospitals; 3988 medical–surgical nurses 4 One year, in-hospital 5 30-day risk-adjusted mortality were from : Discharge Abstract Database (DAD) from the Ontario Ministry of Health and Long-Term Care, and the Ontario Registered Persons Database (RPDB); the patients' socialeconomic status were from the Statistics Canada 1996 Population Data file. Nursing-related variables were from the Ontario Registered Nurse Survey of Hospital Characteristics and the Ontario Hospital Reporting System (OHRS); teaching hospital status was from the Ontario Council of Teaching Hospitals; hospital location (urban or rural) was from the statistics Canada Census 1996 Population Statistical Profiles of Canadian Communities file. From 1 April 1998 to 31 March 1999
Results Quantitative results	A richer skill mix of RNs was found to be associated with lower 30-day mortality, while the dose of nurse staffing was not found to be associated with 30-day mortality. The final multiple regression model showed that a 10% increase in the proportion of RNs across all hospital types was associated with five fewer patient deaths for every 1000 discharged patients. Nursing dose was not included in the final regression model since it was not found a significant result at the first four models. The final regression model adjusted for: years of clinical unit experience for non-urban community hospitals; years of clinical unit experience for teaching hospitals; years of clinical unit experience for urban community hospitals; capacity to work for non-urban community hospitals; capacity to work for teaching hospitals; capacity to work for urban community hospitals.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, sex, 14 categories of pre-existing comorbid conditions, and chronicity of health indicator 2 Adjusted for patients' socialeconomic status; clinical nursing worked hours; availability of role support for nurses; years of experience on the clinical unit; nurse capacity to work; condition of nursing practice environment; continuity of registered nurse care provide; physician expertise 3 Uniform 4 Complete 5 N/A 6 The sample accounted for 4% of all patients discharged from Ontario acute-care hospitals in the study period.

Commentary	It did not adjust for the hospital status and location. The nature of retrospective study could threaten the validity of the results. The model accounted for only 32% of the variance in 30-day risk-adjusted mortality among hospitals; there clearly were other determinants, unknown and unspecified, of 30-day mortality. There is a potential to introduce measurement error, particularly random errors associated with the use of secondary data sources extracted from the Discharge Abstract Database. Little is known about the reliability of OHRS files. Responses in the Ontario Nurse Survey may contain sources of error, no tests of stability were undertaken with the Ontario Nurse Survey and the degree of error in survey responses is unknown.
Research implications	Hospital re-organisation activities that resulted in fewer years of RN experience on their clinical unit contributed to excessive or unnecessary patient mortality. The condition of the nursing environment may be a mediating factor that is itself affected by predictor variables such as nursing skill mix and nurse staffing dose. Replication and refinement of the 30-Day Mortality Model is an important next step in theory development. To test total effects of the predictors, rather than direct effects only, structural equation modelling may be more appropriate.

ID, origin, authors (year)	70, USA, Unruh, L. (2002)
Aims	To examine the changes in licensed nursing staff and to assess the relationship of licensed nursing staff with patient adverse events in hospitals. *Workforce:* Licensed nurses (LN); secondary acute care *Feature:* Number and proportions of LNs – number of LNs; the ratio of LN/patient load, with and without adjusting for patient acuity; the proportion of LNs/total nursing staff *Outcome:* 6 adverse events sensitive to nursing care: atelectasis, decubitus ulcers, falls, pneumonia, post-surgical and treatment infections, urinary tract infections
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 All patients in the hospitals of Pennsylvania 3 211 hospitals yearly, for a total of 1477 during 7 years 4 7 years, in-hospital 5 For nursing personnel and hospital characteristics – the Pennsylvania Department of Health (PDH) and the American Hospital Association (AHA); for adverse events – patient discharge records; for patient-level information – the Pennsylvania Health Care Cost Containment Council (PHC4)
Results Quantitative results	Greater incidence of nearly all adverse events occurred in hospitals with fewer licensed nurses. Regression on the relationship between adverse events, incidence and licensed nurses/total nursing staff. The mean value and percentage change in licensed staff categories was adjusted for patient days of care, and hospital acuity in Pennsylvania Hospitals.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, gender, race, ethnic status, and level of severity upon admission 2 Adjusted for the yearly number of patients, hospital acuity, and other hospitals' characteristics: ownership status, hospital mergers, the number of board-certified physicians, and capacity utilization (occupancy rate/length of stay), a construct of two measures frequently used separately; also adjusted for a year marker that signifies the passage of time from 1991 to 1997, it indicates the influence on adverse events of any changes in the hospital service market or in hospitals that occurred over the time period. 3 Uniform 4 Complete 5 Convenience sampling 6 One state

Commentary	The utilisation of administrative data to ascertain complications during hospitalisation has been controversial. Validity concerns are whether the conditions represented by ICD-9-CM codes were present, or whether the algorithms for ascertaining complications can distinguish between conditions present upon admission, versus those occurring in the hospital. Reliability of data gathered from different sources is a problem. The Pennsylvania records used in this study have been identified as 'intermediate steps in converting clinical observations to an electronic format', and are cleaned up in an effort to produce as reliable data as possible. Poor reporting compliance of hospitals in the early years of the data collection may have reduced the calculated rates of adverse events. Reporting inconsistencies, such as the initial use of only five fields for secondary diagnosis, may have led to the omission of some adverse events in the early years of the study. There are no standard coding procedures among researchers on complications.
Research implications	Need to improve ICD-9-CM coding by the requirement of date markers for all secondary diagnoses so that patient comorbidities can be better distinguished from complications arising from the patient's stay. Given the increases in patient acuity and patient care intensity, when considering licensed staffing targets, a flat licensed nurse/patient ratio, without consideration of patient acuity, may over- or underestimate staffing needs in a particular unit or institution at a given point in time. Therefore, the need to develop or reintroduce flexible staffing systems that take into account daily patient condition severity is essential to both adequate and cost-efficient staffing. Hospitals and policymakers should increase the supply of nurses, including drawing nurses back and attracting more people into nursing. It will be important to bring back both RNs and LPNs in some mix, and to attract new young people to both occupations.

Table A2.10 Volume

ID, origin, authors (year)	75, UK, Bachmann, M.O. *et al.* (2002)
Aims	To evaluate the influence of specialisation on the management and outcome of patients with oesophageal and gastric cancers in National Health Service (NHS) and to examine volume–outcome relationships. *Workforce:* Doctors, secondary *Feature:* Volume; doctor and hosptial *Outcome:* Survival time and operative (30-day) mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective, correlation study 2 Included all patients diagnosed for the first time as having oesophageal or gastric cancer in all 23 acute NHS hospital trusts in the former South and West region of England from July 1996 to June 1997. 3 1512 patients (781 with oesophageal cancer, 731 with gastric cancer) 4 At least 2 years (range 25–41 months, median 31 months) 5 Patient data were extracted from hospital records by three trained researchers. Two outcome measures: operative mortality, defined as death within 30 days of an operation, and risk of death (at least track down to 2 years). Each patient's survival was tracked with the NHS Central Register.
Results Quantitative results	The influence of specialisation on the management and outcomes of patients with oesophageal and gastric cancers was studied. After case mix adjustment, for oesophageal cancer, the operative mortality rate decreased by 40% (odds ratio 0.60, p = 0.047) for each increase of 10 patients in doctors' annual surgical caseloads, and the risk of death decreased by 8% (hazard ratio 0.92, p = 0.021) for each increase of 10 patients in doctors' annual caseloads. For gastric cancer, the operative mortality rate decreased by 41% (odds ratio 0.59) for each increase of 10 patients in doctors' annual surgical caseloads, and the risk of death decreased by 7% (hazard ratio 0.93, p = 0.009) for each increase of 10 patients in hospitals' annual caseloads.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Cox's proportional hazard model was used to adjust for calculating correlation of operative mortality and risk of death with specialisation. For oesophageal and gastric cancer, patients of higher-volume doctors were more likely to have stage I-III disease. Adjustment for case mix and treatment reduced the strength of association of oesophageal cancer patients' survival time with doctor volume, stage and resection, and eliminated the association with hospital volume. For gastric cancer, adjustment increased the strength of the association of survival time with hospital volume and eliminated the association with doctor volume. 2 No 3 There might be slight variation in quality of hospital patient records. Risk of death (survival time) measurement is not uniform because some patients were followed up for 25 months and some for 41 months (median 31 months). 4 Yes. Follow-up completed. 5 Included all patients in the 23 participating hospitals. Assignment of doctors was not random. 6 23 acute NHS hospital trusts in former South and West region of England
Commentary	Large sample size. Limitations were exclusion of patients not admitted to participating hospitals, reliance on hospital sources to identify cases and variable quality of hospital records.
Research implications	Patients of non-specialist doctors and hospitals are less likely to receive effective investigations and treatments.

ID, origin, authors (year)	15, England and Wales, Bachmann, M.O. *et al.* (2003)
Aims	To evaluate the influence of specialisation on the management and outcome of patients with pancreatic cancer *Workforce:* Doctors, secondary *Feature:* Volume; doctor and hosptial *Outcome:* Survival time and operative (30-day) mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective correlation study 2 Included all patients diagnosed for the first time as having pancreatic cancer in any of the 23 acute hospitals in the former South and West National Health Service (NHS) region of England, and in 6 acute hospitals in South Wales from July 96 to June 97 3 782 patients 4 2–3 years 5 Patient data were extracted from hospital records by three trained researchers after notification and again one year after first presentation to hospital. Each patient's survival was tracked with the NHS Central Register for between 25 and 37 months or until death, if earlier. Ecological socioeconomic indicators were provided by Townsend and Carstairs' deprivation scores, which were derived from each patient's postal address.
Results Quantitative results	Patient managed by higher-volume hospitals survived significantly longer (hazard ratio 0.88; p <0.001). They were more likely to undergo cytological examination, resection and biliary stenting. Patients of higher-volume doctors were likely to undergo endoscopic retrograde cholangiopancreatography, percutaneous transhepatic cholangiography, laparoscopy, resection and bypass surgery.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Cox proportional hazards model was used to adjust acuity. 2 For operative mortality, doctor and hospital surgical volumes were used instead of total patient volumes. To convert the adjusted odds ratio and hazard ratios associated with a unit volume difference to the corresponding ratios for larger volume differences, the odds ratios and hazard ratios were exponentiated to the power of the volume differences. 3 Yes 4 Completed. All survivors were followed for at least 2 years (median 31 months). 5 All identified cases provided by the participating hospitals were included. 6 Acute hospitals in the former South and West National Health Service (NHS) region of England, and in 6 acute hospitals in South Wales.
Commentary	Limitations were the exclusion of patients not admitted to participating hospitals, reliance on diverse hospital sources to identify cases, imperfect information on prognostic factors at the time of presentation, the lack of quality-of-life information, which is highly relevant to palliative care.
Research implications	The concentration of pancreatic cancer care into higher-volume hospitals is likely to improve survival even among patients with incurable disease. Priorities for future research are evaluation of the specialisation of cancer care over time, and assessment of the effects of specialisation on patients' quality of life as well as on mortality.

ID, origin, authors (year)	1166, USA, Begg, C.B. *et al.* (2002)
Aims	To examine variations in morbidity after radical prostatectomy for prostate cancer in relation to hospital volume and surgeon volume. *Workforce:* Surgeon, hospital: radical prostatectomy; secondary *Feature:* Surgeon volume – the number of procedures performed by individual surgeons during the study period: low 1–10; medium 11–19; high 20–32; very high 33–121 Hospital volume – the number of procedures performed at individual hospitals during the study period: low 1–33; medium 34–61; high 62–107; very high 114–252 Assuming that 42 percent of patients undergoing prostatectomy are more than 65 years of age, the ranges above corresponded to the total volumes in the analysis are: Surgeon volume: low 1–4; medium 5–9; high 10–15; very high 16–58. Hospital volume: low 1–16; medium 17–28; high 29–50; very high 51–120. *Outcome:* Postoperative death, postoperative complications, late urinary complications, and long-term incontinence.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Excluded patients less than 65 years of age, who were not treated in a Surveillance, Epidemiology and End Results (SEER) state, were not enrolled in both Part A and Part B of Medicare, or were not listed in Medicare records as having undergone prostatectomy within six months after the diagnosis. For study of variations according to surgeon, the cohort was reduced to the 10,737 patients and whose surgeons could be identified in Medicare records. For the 198 patients with more than one surgeon, the study selected the surgeon who had the larger volume of patients for analysis. 3 11,522 patients from 403 hospitals among 999 surgeons in six metropolitan areas and five states 4 4 years; in-hospital 5 Data were from SEER – Medicare linked database to evaluate health-related outcomes after radical prostatectomy; 1992–1996.
Results Quantitative results	Neither hospital volume nor surgeon volume was significantly associated with surgery-related death, but significant reverse relationship was found in volume and postoperative complications and late urinary complications, results for long-term preservation of continence were less clear-cut.

Relation between hospital volume and outcomes (11,522 patients)

Hospitals % of patients	***low (280)***	***medium (67)***	***high (37)***	***v. high (19)***	p-***value****
Surgery-related death					
30 days	0.5	0.5	0.5	0.5	0.92 (0.81)
60 days	0.6	0.6	0.6	0.5	0.94 (0.68)
Postoperative complications	32	31	30	27	0.02 (0.03)
Late urinary complications					
Symptoms or procedures	28	29	23	20	<0.001(<0.001)
Major events	18	19	16	13	<0.001(<0.001)
Long-term incontinence					
Symptoms or procedures	19	19	18	18	0.38 (0.21)
Major events	6.5	6.4	7.0	7.6	0.22 (0.34)

Relation between surgeon volume and outcomes (10,737 patients)

low (642)	***medium (198)***	***high (103)***	***v. high (56)***	p-***value****
0.4	0.5	0.5	0.5	0.71 (0.74)
0.5	0.5	0.6	0.6	0.74 (0.59)
32	31	30	26	0.008 (<0.001
28	26	27	20	0.003 (0.001)
19	18	17	14	0.01(0.01)
20	20	19	16	0.08 (0.04)
7.3	7.2	6.7	6.6	0.82 (0.34)

* *p*-values were adjusted only for within-hospital correlations

** *p*-values were adjusted for both case mix and within-hospital correlations

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, race, stage of disease at diagnosis, and the extent of coexisting illnesses according to the Romano modification of the Charlson index. 2 Adjusted for within hospital or within-surgeon correlations in outcome with use of the generalised-estimating-equations modification of logistic regression. Tested the validity of Medicare claims data on late urinary complications and incontinence by comparing the claim records with the directly observed clinical outcomes through questionnaire in the Prostate Cancer Outcomes Study (PCOS). 3 Uniform 4 4 years 5 No 6 Represent 14% of the population of the USA.
Commentary	No specific definition on the within-hospital or within-surgeon characteristics. The Medicare claims are relatively low-sensitive to detect incontinence, thus may limit the power to detect an effect of hospital or surgeon volume on this outcome. The coexisting conditions and age are a crude and incomplete measure of risk, which could not rule out the possibility that the observed variations may be due to inadequate adjustment for risk factors, especially in the analysis of individual surgeons.
Research implications	An important factor that can influence the quality of surgical care is the availability of ongoing feedback about adverse outcomes to surgical teams and individual surgeons, and the events that are not life-threatening but that affect the patient's quality of life may be less readily apparent and, indeed, may not be observed by surgeons at all. Thus we need more careful scrutiny of adverse outcomes so as to reduce the burden of suffering among patients who undergo surgery for prostate cancer. Need more active educational efforts by professional societies to optimise the quality of surgical care.

ID, origin, authors (year)	1165, USA, Birkmeyer, J.D. *et al.* (2002)
Aims	The importance of hospital volume to the operative mortality associated with six types of cardiovascular procedures and eight types of major cancer. *Workforce:* Hospital: cardiovascular and cancer surgery; secondary care *Feature:* Hospital volume: total number of procedures performed per year: low, very low, medium, high and very high; different volume cut-off point for different procedures *Outcome:* In-hospital mortality or within 30-day mortality after the index procedure.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Only patients covered by fee-for-service arrangement in Medicare records are included; 10% of Medicare patients who were enrolled in risk-bearing health maintenance organisations were excluded; patients who were under 65 years of age or over 99 years of age were excluded. 3 2,500,000 patients from 1086 hospitals 4 5 years; in-hospital or 30 days after the index procedure 5 Patient data were from the Medicare Provider Analysis and Review (MEDPAR) files and the denominator files from the Center for Medicare and Medicaid Services; 1994–1999.
Results Quantitative results	Higher-volume hospitals had lower operative mortality rates for six types of cardiovascular procedures (coronary-artery bypass grafting, heat-valve replacement, carotid endarterectomy, lower-extremity bypass, elective repair of abdominal aortic aneurysm) and eight types (colectomy, gastrectomy, esophagectomy, pancreatic resection, nephrectomy, cystectomy, pulmonary resection, pneumonectomy) of major cancer resections. Dramatic differences in mortality between very low-volume and very high-volume hospitals were observed for pancreatic resection and esophagectomy, whereas relatively small differences in mortality (1% or less) were found for three procedures.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, sex, race, and their interactions; the year of the procedure, the relative urgency of the index admission, the presence of coexisting conditions. 2 Adjusted for the patient mean income from Social Security according to the ZIP codes. 3 Uniform 4 Complete 5 No 6 National
Commentary	In order to avoid selecting cut-off points that could maximise the associations between volume and outcome, the cut-off points were established before mortality was examined, and the points were varied for different procedures. Only studied Medicare patients; the results may not be generalisable to patients under 65 years of age. The volume was estimated total hospital volume by extrapolating from Medicare volume, not by direct measurement, which may remain some degree of misclassification of hospital volume status. The study did not attempt to adjust for characteristics of the provider that are likely to be highly correlated with volume. The administrative data may not have accounted adequately for differences in case mix among strata of hospital volume; data lack the information of patients' severity of illness.
Research implications	Analysis that aimed to assess the independent effect of hospital volume would need to account for other variables that may influence mortality, including hospital size and teaching status, the volume of procedures performed by a particular surgeon, and staffing patterns in the intensive care unit. The study supports the minimal volume standards for different surgeries. Lack of information on the procedure-specific mortality at the level of the individual hospital. In the absence of better information about surgical quality, patients undergoing many types of procedures can substantially improve their odds of survival by selecting a high-volume hospital near them.

ID, origin, authors (year)	173, USA, Brow, P.P. *et al.* (2001)
Aims	High-volume off-pump coronary artery bypass (OPCAB) sites have better clinical outcomes. *Workforce:* Surgical team: OPCAB; secondary care *Feature:* Volume of off-pump coronary artery bypass graft (CABG) procedures: high volume ≥100; low volume < 100 *Outcome:* Mortality, complications: shock–haemorrhage, neurological, cardiac, renal, mechanical, implant infection, postoperative infection, septicaemia, respiratory, pneumonia, peripheral vascular
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional observational study 2 Patients who underwent CABG procedures 3 16,988 consecutive patients in 72 hospital 4 1 year; in-hospital 5 The Healthcare Company case mix database (HCA); 1 January 1999 to 31 December 31 1999
Results Quantitative results	Higher volumes of OPCAB operations would be associated with lower patient and facility complication rates than lower volumes. The data were presented in two ways: *the patient profiles* included patient means, standard deviations, and *p*-values for each of the respective variables in the low- and high-volume OPCAB sites; *the hospital profiles* aggregated patients' data for each hospital and included hospital means, standard deviations, and *p*-values for each of these variables.

	Patient profile: mean (SD)			**Hospital profile: mean (SD)**		
	low volume	***high volume***	**p-*value***	***low volume***	***high volume***	**p-*value***
Patient mortality	2.87 (16.70)	2.85 (16.63)	0.952	3.12 (2.38)	2.96 (1.24)	0.787
Shock–haemorrhage	3.79 (19.09)	2.16 (14.54)	0.000	4.17 (3.77)	2.72 (1.18)	0.044
Neurological	1.45 (11.96)	0.83 (9.10)	0.025	1.13 (1.24)	1.12 (0.88)	0.976
Cardiac	7.47 (26.28)	3.04 (17.18)	0.000	7.46 (6.36)	4.58 (3.68)	0.126
Renal	0.97 (9.80)	0.34 (5.58)	0.005	0.83 (1.18)	0.26 (0.58)	0.068
Mechanical	0.29 (5.36)	0.25 (4.95)	0.720	0.43 (0.73)	0.16 (0.19)	0.034
Implant infection	0.43 (6.53)	0.29 (5.42)	0.309	0.50 (0.89)	0.18 (0.30)	0.069
Postoperative infection	0.99 (9.90)	0.59 (7.65)	0.079	1.50 (3.23)	0.61 (0.39)	0.041
Septicemia	2.22 (14.74)	1.37 (11.64)	0.013	2.23 (1.93)	1.61 (1.81)	0.446
Respiratory	3.30 (17.86)	1.23 (11.01)	0.000	3.47 (4.62)	1.98 (2.50)	0.236
Pneumonia	0.91 (9.50)	1.13 (10.57)	0.375	0.95 (1.22)	1.17 (0.51)	0.402
Peripheral vascular	0.50 (5.04)	0.15 (3.84)	0.257	0.31 (0.67)	0.18 (0.28)	0.357

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, sex, smoking, history of tobacco use, comorbid conditions: chronic obstructive pulmonary disease, insulin-dependent diabetes, noninsulin-dependent diabetes, acute renal failure, acute renal failure, chronic renal failure, unspecified renal failure, cardioshock shock, hypertension, acute myocardial infarction, old myocardial infarction, cardiomyopathy, congestive heart failure, peripheral vascular disease, and endocarditis in logistic regression. 2 No 3 Uniform 4 Complete 5 N/A 6 National

Commentary	The HCA case mix database is an administrative database and lacks clinical details that would be useful in segmenting patients and clinical characteristics. The timing of events is not known (preoperatively, intraoperatively, or postoperatively). The physician's intention to treat could not be identified. The study used the facility's number of off-pump procedures, not the individual surgeon's off-pump experience. It is possible that the individual surgeon's experience could be of more importance than that of the overall surgical team.
Research implications	More studies needed on the impact of individual surgeon's volume on the patients' outcome.

ID, origin, authors (year)	676, USA, Finlayson, E.V.A. and Birkmeyer, J.D. (2003)
Aims	To evaluate the impact of surgeon and hospital characteristics on patient outcomes in colorectal surgery *Workforce*: Surgeon and hospital: colorectal surgery; secondary care *Feature:* Surgeon volume, hospital volume, and surgeon board certification *Outcomes*: Surgical and late mortality, and rate of recurrence.
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Literature review 2 Include: studies examined the association between provider characteristic, e.g. surgeon volume, and patient outcomes in colorectal surgery 3 Not reported 4 Not reported 5 Not reported 6 Studies varied widely in what volume cut-off points were used to define low- and high-volume providers. Many studies examined colon and rectal procedures in aggregate, obscuring the possibility that volume may play a different role in the outcomes with each surgery. Many studies did not account for the interaction between surgeon and hospital characteristics. Studies focused on different populations, some included aged 65 or older whereas others include all ages. Many of the large studies were based on administrative data, which may lack the clinical detail to fully account for differences in case-mix across providers. 7 A narrative review described the findings of the individual articles
Results Quantitative results	For colon cancer, preponderance of evidence suggests that patients undergoing colon resection at high-column hospitals have small but clinically meaningful reductions in surgical and late mortality. Surgical and late mortality reductions also have been documented for high-volume surgeons. For rectal cancer surgery, numerous studies suggest that surgeons with more experience and colorectal subspecialty training have better results, including lower rates of local recurrence and late mortality.
Commentary	It is a literature review, no systematic methods in reviewing the studies.
Research implications	The role of volume in surgical morbidity has not been characterized. The association between hospital and surgeon variables and resource use is not well understood.

ID, origin, authors (year)	610, USA, Hillner, B.E., Smith, T.J. and Desch, C.E. (2000)
Aims	To search for evidence that hospital or physician volume or specialty affects the outcomes of cancer care. *Workforce*: Physician and hospital: cancer care; secondary care *Feature*: Physician volume, hospital volume, physician specialisation, and hospital specialisation *Outcomes*: In-hospital mortality or 30-day mortality
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Literature review 2 Inclusion: studies focused on the relationship of volume and health outcome of cancer care; studies that stratified or adjusted for clinical stage. Exclusion: reports related to screening or early detection and surveys of physician attitudes or practices based on hypothetical patients. The bibliography of each article was reviewed for other potentially relevant citations. 3 Not reported 4 Not reported 5 Sources searched: Medline (1988–1999) 6 All reports used retrospective data and predominantly used data collected in the 1980s. All reports were stratified by or adjusted for clinical stage, however, the other risk adjustment were not reported for every report. Some reports had the motivation that citing an association of higher volume with better outcome could strengthen their conclusions, which may cause publication bias. Almost all reports used convenience samples and had no pre-planned statistical power or effect-size estimates. Comorbidity was usually inferred from administrative claims, not from specific clinical indices or databases; therefore, the ability to adequately control for case mix is weaker. The vast majority of studies focused on the short-term outcomes of cancers for which the primary mode of therapy was surgery performed with curative intent. Long-term outcomes of these surgical therapies were substantially fewer. Some of the reports used un-uniformed data for comparison. 7 A narrative review described the findings of the individual reports. Not every study was reported regarding whether it identified and controlled for case mix by adjusting for demographics and/or comorbidity; however, it seems that the studies that included adjustments for comorbidity were specifically highlighted. All studies were classified by procedures; some of them had detailed description on data sources, units of analysis, country, and risk adjustment. But no summary according to the quality of each primary study.
Results Quantitative results	Most reports support a positive volume-outcome relationship in initial cancer treatment. For cancers treated with technologically complex surgical procedures (non-small-cell lung cancer, pancreatic, oesophageal, and gastric cancer), an extensive, consistent literature supported a volume-outcome relationship. For the cancer primarily treated with low-risk surgery (colon, breast, prostate, and ovarian cancer), there were few studies; an association with hospital and surgeon volume in colon cancer varied with the volume threshold. For breast cancer, British studies found that physician specialty and volume were associated with improved long-term outcomes, and the single American report showed an association between hospital volume of initial surgery and better 5-year survival. For nonsurgical cancers (lymphomas, testicular cancer, leukaemia), there were few studies but all consistently showed better long-term outcomes associated with higher hospital volume or specialty focus. For recurrent or metastatic cancer, there was no study.
Commentary	This review is not a systematic review and it did not pool the quantitative data. Since this review is not solely on volume–outcome relationship, some of the primary studies included only investigated the specialisation of the workforce. There is a lack of clear summary of results according to the quality of the studies for individual procedures.

Research implications	The well-defined first identification, and the tumor-node-metastasis taxonomy, actual cancer care should and can be prospectively measured, assessed, and benchmarked. For all forms of cancer, efforts to concentrate its initial care would be appropriate. Long-term outcomes of the surgical therapies were substantially fewer. The specific processes and hospital/organisational factors that lead to or that are associated with the superior outcomes in specific hospitals or physician specialities have not been deciphered. One process area that may account for better outcomes is the reorganising of care from diversified locations into a single-site multidisciplinary clinic, but this benefit did not have pre-/post-treatment comparative studies. No study examined broader outcomes, such as level of pain control of patient, family satisfaction, related to hospital or physician characteristics. It is difficult to determine the direction of the causal relationship: whether volume affects quality or whether better units and clinicians attract more patients.

ID, origin, authors (year)	569, USA, Katz, J.N. *et al.* (2001)
Aims	To assess the relationship between surgeon and hospital procedure volume and mortality and complications in the first 90 days postoperatively in primary and revision total hip replacement. *Workforce:* Surgeon and hospital: total hip replacement (THR); secondary care *Feature:* Hospital volume: 1 Annual volume of primary THR: very low 1–10; low 11–25; medium 26–50; high 51–100; very high >100 2 Annual volume of revision THR: very low 1–5; low 6–10; medium 11–25; high 25–50; very high >50 Surgeon volume: 1 Annual volume of primary THR: very low 1–5; low 6-10; medium 11–25; high 25–50; very high >50 2 Annual volume of revision THR: low 1–3; medium 4–10; high >10 O*utcome:* Mortality, dislocation, deep inflection, and pulmonary embolus in the first 90 days postoperatively
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 The study included patients who had had elective primary or revision total hip replacement and excluded patients who were less than 65 years old and those with codes indicating infection of the hip, metastatic or bone cancer, conversion of hemiarthroplasty (or other hip surgery) to total hip replacement, or fracture of the hip or femur. For the analysis of revision total hip replacement, it excluded patients enrolled in a health maintenance organisation, patients who were not enrolled in both parts of Medicare and those who were not residents of the USA. 3 59,521 patients for elective primary total hip replacement; 12,956 patients for revision total hip replacement 4 1 year 5 Patients' data were from Medicare claims data; surgeons data were from the Medicare Unique Physician Identification Number (UPIN); hospitals' data were from the 1995 American Hospitals Association Survey; 1 July 1995 to 30 June 1996.
Results Quantitative results	Higher hospital volume was significantly associated with lower rates of mortality and dislocation after primary total hip replacement; higher surgeon volume was significantly associated with a lower rate of dislocation, and less strongly, with a lower rate of deep hip infection, the revision THR had similar results, the only exception is that surgeon volume, but not hospital volume, was associated with mortality. *Association between hospital and surgeon procedure volumes and select outcome of primary total hip replacement* **Hospital volume – rate of outcome/adjusted OR (95% CI)** ***v. low*** / ***low*** / ***medium*** / ***high*** / ***v. high*** / ***v. high*** Mortality: 1.3%/1.0 / 1.0%/0.82 (0.62, 1.07) / 0.9%/0.72 (0.54, 0.95) / 0.9%/0.68 (0.51, 0.92) / 0.7%/0.58 (0.38, 0.89) / 0.7%/0.34 (0.95, 1.62) Dislocation: 4.4%/1.0 / 3.8%/0.96 (0.82, 1.17) / 2.9%/0.79 (0.67, 0.93) / 2.5%/0.72 (0.60, 0.87) / 2.2%/0.77 (0.58, 1.03) / 1.5%/0.49 (0.34, 0.69) Deep infection: 0.4%/1.0 / 0.3%/.84 (0.52, 1.37) / 0.2%/.56 (0.33, 0.96) / 0.2%/.74 (0.42, 1.32) / 0.1%/0.52 (0.22, 1.22) / 0.1%/0.28 (0.07, 1.11) Pulmonary embolus: 1.1%/1.0 / 1.0%/0.86 (0.64, 1.15) / 1.0%/0.89 (0.66, 1.21) / 0.8%/0.83 (0.60, 1.14) / 0.8%/0.79 (0.51, 1.23) / 0.7%/0.73 (0.44, 1.21)

Surgeon volume – rate of outcome / adjusted OR (95% CI)

v. low	*low*	*medium*	*high*
1.1%/1.0	1.0%/0.98 (0.78, 1.23)	0.9%/0.97 (0.77, 1.22)	0.8%/1.10 (0.95, 1.54)
4.2%/1.0	3.4%/0.85 (0.76, 0.96)	2.6%/0.68 (0.59, 0.78)	2.4%/0.68 (0.54, 0.86)
0.3%/1.0	0.3%/0.9 (0.59, 1.37)	0.2%/0.80 (0.51, 1.26)	0.1%/0.64 (0.30, 1.36)
1.0%/1.0	1.0%/0.98 (0.78, 1.23)	0.9%/0.91 (0.72, 1.14)	0.7%/0.75 (0.51, 1.08)

Risk of dislocation associated with surgeon volume of primary total hip replacement, stratified by hospital volume

Hospital volume – rate of outcome / adjusted OR (95% CI)

	medium	*high*	*very high*
Surgeon volume			
Very low	3.7%/1.0	3.2%/1.0	2.5%/1.0
Low	3.0%/0.83 (0.66, 1.05)	3.4%/1.1 (0.81, 1.45)	2.5%/1.2 (0.57, 2.45)
Medium	2.5%/0.69 (0.54, 0.89)	2.2%/.70 (0.52, 0.95)	2.6%/1.2 (0.61, 2.20)
High	2.9%/0.84 (0.53, 1.31)	2.1%/.65 (0.46, 0.92)	2.6%/1.2 (0.68, 2.20)
Very high	1.3%/0.34 (0.10, 1.13)	1.1%/.33 (0.19, .59)	1.7%/.95 (0.51, 1.77)

Association between hospital and surgeon procedure volumes and select outcomes of revision total hip replacement

	Hospital volume – rate of outcome / adjusted OR (95% CI)				**Surgeon volume – rate of outcome / adjusted OR (95% CI)**			
	v. low	***low***	***medium***	***high***	***v. high***	***low***	***medium***	***high***
Mortality	3.5%/1.0	2.6%/0.85 (0.62, 1.15)	2.1%/0.74 (0.54, 1.00)	1.5%/0.67 (0.40, 1.11)	1.8%/0.85 (0.43, 1.67)	3.1%/1.0	2.2%/0.78 (0.59, 1.03)	1.5%/0.65 (0.44, 0.96)
Dislocation	9.8%/1.0	8.6%/0.90 (0.75, 1.08)	8.4%/0.90 (0.75, 1.09)	7.0%/0.75 (0.56, 1.02)	4.2%/0.45 (0.30, 0.66)	9.1%/1.0	8.7%/1.04 (0.89, 1.21)	6.1%/.84 (0.67, 1.06)
Deep infection	0.9%/1.0	1.1%/1.31 (0.78, 2.21)	1.0%/1.39 (0.84, .2.31)	0.9%/1.36 (0.64, 2.92)	0.5%/0.78 (0.29, 2.10)	1.0%/1.0	1.0%/0.97 (0.61, 1.55)	0.7%/0.64 (0.33, 1.24)
Pulmonary embolus	0.7%/1.0	1.1%/1.63 (0.94, 2.81)	0.7%/1.01 (0.54, 1.90)	0.5%/.67 (0.29, 1.57)	0.7%/0.91 (0.40, 2.06)	0.7%/1.0	1.0%/1.44 (0.89, 2.34)	0.6%/1.00 (0.53, 1.90)

Quality appraisal

1 Case mix adjustment
2 Other adjustment
3 Uniform data collection
4 Participant follow-up
5 Random sampling
6 Geographical dispersal

1 Adjusted for patients' age, gender, race, Medicaid eligibility (a surrogate for low income), arthritic diagnosis, and comorbidity index.
2 Adjusted for urban and rural location of the hospital, the ratio of nurses to discharges, and the teaching and ownership status of the hospital; adjusted for the number of years since the surgeon graduated from medical school. The hospital volume models were adjusted for surgeon volume, and the surgeon–volume models were adjusted for hospital volume.
3 Uniform
4 Ninety days after operation
5 No
6 National

Commentary	The findings for revision total hip replacement are less precise because of a smaller sample size. The analyses failed to reveal discrete volume thresholds that distinguished favourable from poor outcomes, but rather they showed a steady trend across all volume strata toward better outcomes associated with higher volume. Key factors such as the complexity of the surgery and preoperative and postoperative psychological and physical functional status and pain are not captured in claims data. Exclusion of Medicare patients who belonged to a health maintenance organisation may have limited generalisability slightly.
Research implications	The trade-off between the comfort of having surgery at a community centre and the better outcomes in referral centres should be examined explicitly. The effects of procedure volume on pain relief, functional improvement, and durability of the implant should be examined to provide a more complete picture of the influence of volume on outcome. Research is needed to identify the aspects of the processes of care and the care setting that provide better outcomes. It would be preferable to urge all centres to adapt these features than to simply close low-volume centres. Regionalisation may be difficult in areas where some patients might be unable to travel to the referral centre.

ID, origin, authors (year)	679, USA, Katz, J.N. *et al.* (2002)
Aims	To evaluate whether hospital volume and surgeon volume of total hip replacements (THRs) are associated with patient-reported functional status and satisfaction with surgery 3 years postoperatively. *Workforce:* Hospital and surgeon; secondary care *Feature:* Hospital volume: the total number of THRs performed per year: 1–12, 13–100, >100 (primary THR cohort); 1–30, 31–100, >100 (revision THR cohort) Surgeon volume: the total number of THRs performed per year: ≤12, >12 *Outcome:* Patient-reported functional status and satisfaction with surgery 3 years after the surgery: measured by the Harris hip score (0–100 from the worst to the best, the lowest 10% Harris hip score corresponding to <55 in the primary THR cohort and <40 in the revision THR cohort) and the satisfaction score (0–100 from the worst to the best; scores <50 indicate that the patient was dissatisfied).
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Population-based retrospective cohort study. 2 Include patients aged 65 years or older, who underwent elective primary or revision THR. Excluded patients with infection of the hip, metastatic cancer or bone cancer, conversion of hemiarthroplasty (or other hip surgery) to THR, and (for primary THR) fracture of the hip or femur. 3 For primary hip replacement, a cohort of 958 patients returned completed questionnaires, 926 patients among them were analysed, since they had complete data on hospital and surgeon volume. 4 3 years post-operation 5 Medicare claims: provided data on the volume of primary and revision THRs performed in 1995 by the surgeon and in the hospital; also provided data on age, sex, and eligibility for Medicaid (a surrogate for low income). Medical record: provided preoperative weight and height, underlying arthritis diagnosis, medical comorbidities, whether cement was used to implant the femoral and acetabular prosthetic components, the approach, and whether the patient had previously undergone hip, knee, or spine surgery. For revision surgery, the medical record was also scrutinized to ascertain whether there had been prior revisions of the index hip, whether a bone graft was used, and if so, whether structural allograft was used. Survey: questionnaires included several validated measures of pain and functional status. The measurement scales included the WOMAC pain scores, Harris hip score and satisfaction scores, and were all converted into 0-100 score, with 100 representing the best possible score; 1995–1998.
Results Quantitative results	Hospital and surgeon volume had little effect on functional status 3 years following primary and revision THR, after adjusting for socioeconomic and clinical variables. However, satisfaction with primary THR is greater among patients who underwent surgery in high-volume centres, and satisfaction with revisions is greater among patients whose operations were performed by higher-volume surgeons.

Association between volume and Harris hip score (95% CI) Lowest 10% Harris hip score

	primary THR		***revision THR***	
	Crude OR	Adjusted OR	Crude OR	Adjusted OR
Hospital volume				
Low	1.78 (0.90–3.54)	1.29 (0.64–2.62)	1.83 (1.08–3.11)	0.90 (0.40–1.99)
Medium	1.32 (0.70–2.49)	1.14 (0.63–2.06)	1.22 (0.73–2.03)	0.94 (0.45–1.95)
High	1.00	1.00	1.00	1.00
Surgeon volume				
≤12	1.26 (0.81–1.99)	–	1.63 (0.93–2.85)	1.45 (0.80–2.96)

Association between volume and dissatisfaction (95% CI) Satisfaction score < 50

	primary THR		***revision THR***	
	Crude OR	Adjusted OR	Crude OR	Adjusted OR
Hospital volume				
Low	2.15 (1.20–3.85)	2.06 (1.15–3.69)	1.26 (0.80–1.97)	0.81 (0.44–1.48)
Medium	1.29 (0.74–2.26)	1.22 (0.70–1.13)	1.12 (0.77–1.63)	0.85 (0.54-1.33)
High	1.00	1.00	1.00	1.00
Surgeon volume				
≤12	1.07 (0.68–1.68)	–	1.68 (1.14–2.46)	1.77 (1.11–2.82)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patient preoperative functional status, prior hip, knee, and spine surgery. 2 Adjusted for patient income, education. 3 Uniform 4 Retrieve the data. 5 Stratified random sample 6 3 states
Commentary	The study had several strengths: the sample was population-based; outcomes were assessed with standardised, previously validated scales; the research team was not involved in the care of patients, which precluded observer bias; the analyses accounted for potential confounding by patient characteristics and were simultaneously adjusted for hospital and surgeon volume and for clustering of patients within hospitals. The study also has important limitations: the response rates were 51% of all eligible patients who underwent primary THR and 39% of patient who underwent revision THR; it used the Medicare data, setting a limit on age of and over 65 years; it did not use a prospective preoperative function assessment – instead, it used variables that are recalled with moderate accuracy; the crude measure of surgical complexity did not capture many of the more subtle aspects of the complexity of revision THR; prosthesis failure leading to revision occurs rarely at 3 years, preventing evaluation of this important and costly outcome; there is a reporting problem with the patient satisfaction; it is hard to recall accurately the satisfaction with the surgery 3 years later.
Research implications	The study suggested that a regionalisation policy would have little additional benefit after the perioperative period. It also showed that older, less educated, poorer, and more functionally disabled patients would be disproportionately affected by a policy that shifts patients out of low-volume hospitals. It argued against a blanket regionalisation policy and suggested that the potential trade-offs between having THR in a small-volume or a large-volume centre should be evaluated explicitly in decision analytic models. It provided a template for research on other procedures such as cardiac and cancer surgery to determine whether the short-term advantages of high-volume hospitals documented for these procedures persist over a longer period and extend to a broader set of outcomes. The analyses should be extended in the future to younger patients undergoing THR. Longer-term follow-up of the revision THR cohort is critical.

ID, origin, authors (year)	79, Canada, Klein, M.C. *et al.* (2002)
Aims	To determine if the practice–volume relations exist in maternity care practice by family doctors *Workforce:* Family physicians – maternity care; primary care *Feature:* Physician volume – the number of births attended each year: <12, 12–24, >25 *Outcome:* Maternal morbidity, 5-minute Apgar score and admission of the baby to the neonatal intensive care unit or special care unit (NICU or SCU)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Cross-sectional observational study 2 All births excluding multiple gestations 3 152 family physicians who attended a total of 4444 singleton births 4 6 months; in-hospital 5 Data on all births are collected from the medical records of BC Women's Hospital and Health Centre; data on ethnicity are abstracted through a structured nursing form, a structured antenatal form and physician progress notes; April 1997 to August 1998.
Results Quantitative results	Family physicians' delivery volumes were not associated with adverse outcomes for mothers or newborns. *Maternal and new born outcomes and physician's delivery volume; no. (and %) of mothers or newborns (n= 4444)* **Total** (n=4444) / **Low** (n=549) / **Medium** (n=871) / **High** (n=3024) / ***p*-value** Complex maternal morbidity: 746 (16.8) / 92 (16.8) / 164 (18.8) / 490 (16.2) / 0.189 5-min Apgar score <7: 158 (3.6) / 22 (4.0) / 25 (2.9) / 111 (3.7) / 0.441 Admission to NICU or SCU: 507 (11.4) / 64 (11.6) / 101 (11.6) / 342 (11.3) / 0.954 *Multivariate odds ratios for association between physician's delivery volume over 18-month study period and maternal and newborn outcomes (n=4267)* Adjusted OR (95% CI / *p*-value) ***low (reference)*** / ***medium*** / ***high*** Complex maternal morbidity: 1.0 (–/–) / 1.137 (0.845–1.529/ 0.398) / 0.960 (0.743–1.242/0.758) 5-min Apgar score <7: 1.0 (–/–) / 0.652 (0.339–1.251/ 0.198) / 0.908 (0.540–1.524/0.741) Admission to NICU or SCU: 1.0 (–/–) / 0.862 (0.584–1.274/ 0.457) / 0.849 (0.610–1.181/0.332)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Adjusted for parity, pregnancy-induced hypertension, gestational diabetes, ethnicity, lone parent status, maternal age, gestational age, birthweight and head circumference at birth in stepwise logistic regression. 2 No 3 Uniform 4 Data for 177 births were missing in the multivariate analysis. 5 No 6 Only one hospital

Commentary	The study may include healthier patients than might be expected. The study did not include the number of years in practice as a variable, which might affect maternal and newborn outcomes. The results cannot be generalised to smaller centres, where obstetric and paediatric consults are not as readily available.
Research implications	Further work in other institutions, with similar data collection methods and adjustment for case mix, risk, hospital size, and urban or rural location as well as number of years of experience of the physician, is needed to validate the findings. The result of this study might revise the SOGC policy and provincial guidelines which recommended that 'physicians with low volumes of obstetrical patients should restrict their practice to "normal" obstetrics and should update their skills every 2 to 3 years'.

ID, origin, authors (year)	208, USA, Margulies, D.R. *et al.* (2000)
Aims	To test the hypothesis that high volume of patients with Injury Severity Score (ISS) >15 per individual trauma surgeon is associated with improved outcome. *Workforce:* Surgeon: Level I trauma centres; secondary care *Feature:* Surgeon volume: number of patients treated during the study period per surgeon: 0–10, 11–20, 21–35, 36–50, 51–100, >100 *Outcome:* Mortality, intensive care unit length of stay (ICU LOS) and hospital LOS
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Include the patients with ISS >15 3 1754 patients in 5 Los Angeles County adult Level I trauma centres 4 14 months; in centres. 5 Data were obtained from the Department of Health Services – Emergency Medical Services trauma registry; 1 January 1 1998 to 31 March 31 1999
Results Quantitative results	There is no correlation between physician volume and health outcomes in trauma care. When mortality was compared for surgeons with a caseload of fewer than 35/year with those caring for more than 35/year, no difference could be demonstrated (logistic regression analysis: $p = 0.73$). There was also no correlation between per-surgeon caseload and ICU LOS ($r = 0.09$) and hospital LOS ($r = 0.03$); however, the paper did not report the regression analysis for LOS and surgeon caseload. *Logistic regression analyses for per-surgeon case-load of mortality* **Mortality** **SE** ***df*** ***p*-value** ***r*** Per-surgeon caseload 0.002 0.003 1 0.438 0.000
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, gender, systolic blood pressure (SBP) at admission, Coma Scale Score, Injury Severity Score, Probability of survival, mechanism of injury (blunt or penetrating), need or no need for laparotomy, and with or without head injury. 2 Institutional volume 3 Uniformed 4 Complete 5 No 6 Only represents well-established urban Level I trauma centre; it does not apply to other centres in which situations may differ.
Commentary	The study did not report in detail the regression analysis for the relationship between surgeon caseload and ICU/hospital LOS, nor did it report in detail the regression analysis of mortality and surgeon caseload when it did the comparison between the surgeons with caseload >35/year and the surgeons with caseload <35/year. The mature trauma system itself, which can be expected to lower mortality, might offset the effects of caseload on outcomes. There were very few providers who saw very high volume, and it certainly was not better than the lower-volume providers. But that in itself might mean that high-volume providers are more skilful.
Research implications	Need to know moreabout the critical care volume that the surgeon is taking care of, and how that impacts on patient outcome. Need to know more about the operative caseload of the surgeon separate from trauma patients, and how that impacts on outcome. Need to look at quality of care; volume is a surrogate indicator of those measures.

ID, origin, authors (year)	600, UK, Shackley, P., Slack, R., Booth, A. *et al*. (2000)
Aims	To examine the evidence for the existence of a positive volume–outcome relationship in the area of peripheral vascular surgery *Workforce:* Surgeon and hospital: peripheral vascular surgery; secondary care *Feature:* Volume (annual caseload for physician and hospital) *Outcome:* Health outcomes (mortality, morbidity)
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 Exclusion: not written in English; not published in peer-reviewed journal, editorial, letter or abstract; the article did not address the issue of volume and outcome in peripheral vascular surgery. 3 36 cohort studies in total: carotid endarterectomy (CE) (17 retrospective); abdominal aortic aneurysm (AAA) repair (16: 14 retrospective, 2 prospective); other vascular interventions (4, design not reported). Note: one study considered both CE and AAA. 4 Not stated 5 Trials Register of the Peripheral Vascular Disorders Review Group of the Cochrane Collaboration. Electronic databases: Medline, Embase, Science Citation Index, Healthstar, DHSS-data, Helmis and Cochrane Library. Searched over the period 1986–1998. Based on that used by the NHS Centre for Reviews and Dissemination, volume and outcome study (1997). Citation search for seminal articles. 6 Adjustment for case mix categorised into three groups: full adjustment (demographic factors, co-morbidity and severity/stage of illness). Severity/stage of illness reported if separately identified asymptomatic and transient ischemic attacks and amaurosis), partial adjustment (demographics and co-morbidity) and no adjustment. 7 Findings from the 36 studies were combined using narrative alone.
Results Quantitative results	CE: positive volume–outcome for mortality and stroke at physician level. Less support for mortality at hospital level, and no evidence of benefits for stroke. When considering only studies with full adjustment for case mix there is no evidence for positive/negative relationship between volume and outcome. Unruptured AAA: positive volume relationship at both physician and hospital level, particularly strong at hospital level. Ruptured AAA: no evidence of positive volume–outcome relationship. Other vascular interventions: insufficient studies to draw meaningful conclusions.
Commentary	Two independent reviewers, third to resolve discrepancies. Only published studies included. Definitions as to what constitutes high and low volume varied among CE and AAA studies. No pooling of results. Each study reported separately. Mortality is the principal measure of outcome, generally referring to inpatient stay. No sensitivity analyses. Large variation between studies in the adjustment for case mix. Mainly retrospective cohort studies of poor quality. There may be differences in experience/skill of physicians performing un-/ruptured AAA, hence affecting outcome.
Research implications	Prospective cohort studies investigating volume–outcome relationship, with full adjustment for case mix, is needed. Studies using mortality as an outcome measure should follow up participants after inpatient stay. Need for outcome data on quality of life.

ID, origin, authors (year)	63, USA, Solomon, D.H. *et al.* (2002)
Aims	Whether specific hospital-level factors accounted for the association between the volume of total hip replacement (THR) and the 90-day rate of orthopaedic adverse events, defined as dislocation of the prosthetic hip and deep wound infection of the hip. *Workforce:* Surgeons and hospitals: THR surgery; secondary care *Feature:* Volume: number of elective primary THR procedures performed in one study year: • Physician volume: 1–5, 6–10, 11–25, 26–50, >50 (<10, ≥10 for combined hospital–physician–volume analysis) • Hospital volume: 1–25, 25–100, >100 *Outcome:* Dislocation of the prosthetic hip and deep wound infection of the hip.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Include patients who underwent primary THR surgery, exclude those younger than age 65 years, or had evidence of any of the following clinical conditions: hip infection, metastatic cancer or bone cancer, any prior hip surgery, and fracture of the hip or femur. 3 5211 Medicare patients in 167 hospitals in Colorado, Pennsylvania, and Ohio. 4 1 year; in-hospital 5 Medicare Part A and Part B claims, the American Board of Medical Specialties, a hospital survey regarding institution-specific characteristics and structural aspects of the care setting, and the American Hospital Association 1995 Annual Survey; 1995–1996.
Results Quantitative results	Hospital volume was no longer significantly associated with adverse events when the hospital volume was added to the fully adjusted model; therefore the volume of THRs performed by individual surgeons is the most important determinant of orthopaedic complications and should be considered in efforts to improve THR outcomes.

		% of patients with adverse event (n=136)	**Adjusted OR (95% CI)**	***p*-value**	**Unadjusted OR (95% CI)**	***p*-value**
Hospital volume						
1–25		3.6	1.00		1.00	
25–100		2.8	0.84 (0.57–1.25)	0.40	0.69 (0.46–1.04)	0.078
>100		1.4	0.44 (0.28–0.70)	<0.001	0.31 (0.17–0.56)	<0.001
Surgeon volume						
1–5		6.2			1.00	
6–10		3.7			0.58 (0.36–0.92)	0.021
11–25		2.0			0.31 (0.20–0.48)	<0.0001
26–50		1.6			0.24 (0.12–0.49)	<0.0001
>50		0.9			0.14 (0.08–0.24)	<0.0001
Hospital and surgeon volume						
1–25	<10		1.0			
	≥10		0.38 (0.12–0.8)	0.0013		
25–100	<10		1.11 (0.45–1.8)			
	≥10		0.52 (0.31–0.81)	0.00052		
>100	<10		1.11 (0.41–2.7)			
	≥10		0.27 (0.1–0.52)	<0.0001		

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' age, sex, race, Medicaid-eligible, comorbid conditions, rheumatoid arthritis, aseptic necrosis, year of THR. 2 Adjusted for hospital characteristics: urban setting, privately owned, fully accredited by JCAHO, teaching hospital, affiliated with a medical school, training programmes, surgical and other facilities, nursing unit, physical therapy, health care services, social work available on weekends; and surgeon characteristics – certified in orthopaedic surgery. 3 Uniform 4 Complete 5 No 6 Three states
Commentary	The study did not report the adjusted OR for the surgeon volume, and also did not have the specific adjusted OR for hospital and surgeon volume except a graph. The cross-sectional nature of the data precludes determination of whether the association between physician volume and outcomes is causal. The hospital survey has not been validated and was conducted 4 years after the period of study; there might be a response bias. The hospital survey provided hospital-level data that may not apply to all patients; misclassification the patient-related hospital characteristics would bias the findings toward the null. The study focused on a narrow set of outcomes. The study was not generalisable to other US states or other parts of the world.
Research implications	Further investigation needed for the effects of limiting performance of THRs to high-volume providers, and identification of practical strategies for providing patients and physicians with better information with which to make referral decisions. The results of the study implied that surgeon volume (primarily) and hospital volumes (secondarily) are the best indicators of future orthopaedic adverse events in patients undergoing THR surgery; however, it is not certain that increasing a surgeon's (or a hospital's) volume of THRs would actually improve outcomes.

ID, origin, authors (year)	554, Canada, Tu, J.V., Austin, P.C. and Chan, B.T.B. (2001)
Aims	To evaluate the relationship between the average annual volume of cases treated by admitting physicians and mortality after acute myocardial infarction (AMI) *Workforce:* Physician: AMI treatment; secondary care *Feature:* Physician volume: the average annual volume of AMI treated by each physician was calculated by dividing the number of AMI cases treated during the 6-year study period by the number of years the physician actually treated 1 or more AMI patient: 1–5, 6–13, 14v24, >24. *Outcome:* Mortality risk rates for 30 days and 1 year post-AMI
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Included patients admitted to Ontario hospitals with AMI. Excluded the patients if they were younger than 20 years or older than 105 years, were not Ontario residents or had an invalid Ontario health care number; were admitted as transfers from another acute care institution or to a noncardiac surgical service; had an AMI coded as an in-hospital complication; were discharged alive with a length of stay of less than 3 days; or were admitted with an AMI in the year before the index admission. 3 98,194 patients treated by 5374 physicians 4 6 years; in-hospital 5 Patients' data were from the Ontario Myocardial Infarction Database (OMID); admitting physicians were identified by linking the OMID cohort to the Ontario Health Insurance Plan (OHIP); additional characteristics of the physicians were from Corporate Providers Database of the Ontario Ministry of Health and the Southam Medical Database; 1 April 1992 to 31 March 1998.
Results Quantitative results	There is a strong inverse association between average annual volume of AMI cases treated by admitting physicians and patient mortality after an AMI.

Physician volume quartile, AMI cases per year

	1–5	***6–13***	***14–24***	***>24***
Compared with high-volume physicians within same specialty				
Cardiologists	1.48 (1.17–1.86)	1.22 (1.06–1.41)	1.09 (0.97–1.23)	Referent
General internists	1.57 (1.30–1.90)	1.27 (1.06–1.51)	1.25 (1.05–1.48)	Referent
Family physicians	1.32 (0.98–1.78)	1.26 (0.92–1.72)	Referent	N/A
Other	2.49 (1.92–3.24)	1.00 (0.77–1.31)	1.09 (0.82–1.44)	Referent
Compared with cardiologists within same volume category				
Cardiologists	Referent	Referent	Referent	Referent
General internists	1.02 (0.80–1.30)	0.99 (0.85–1.15)	1.09 (0.96–1.24)	0.95 (0.81–1.31)
Family physicians	0.78 (0.62–0.98)	0.90 (0.77–1.06)	0.88 (0.66–1.16)	N/A
Other	1.98 (1.56–2.51)	0.96 (0.82–1.14)	1.17 (0.99–1.38)	1.17 (0.91–1.51)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients' characteristics (age, sex, predicted 30-day mortality, socioeconomic status). 2 Adjusted for other physicians' characteristics (specialty, age, sex); hospital characteristics (hospital volume and teaching status, availability of on-site revascularisation facilities) 3 Uniform 4 Complete 5 No 6 One state

Commentary	It was not possible to adjust for all possible clinical factors that influence mortality after an AMI by using administrative databases. The undercoding of comorbid conditions in the administrative database may have reduced the ability of the statistical regression models to adjust for factors that may affect the physician volume–outcome relationship. The study did not have information on in-hospital use of various therapies such as thrombolytics, which could partially explain the relationship.
Research implications	High-volume physicians may be better at recognising an AMI and interpreting difficult electrocardiograms. They may be faster at making decisions regarding thrombolytics, choose more appropriate risk stratification tests, make more appropriate referral decisions, and be more skilled at treating complications. These possible explanations will need to be investigated in future studies. The exact mechanisms contributing to this complex phenomenon remain to be elucidated. Both shifting the care of AMI patients to high-volume physicians and developing strategies to improve the clinical expertise of low-volume physicians may lead to better patient outcomes.

ID, origin, authors (year)	604, USA, Tulford, J.M. *et al.* (2000)
Aims	To investigate whether an increase in patient volume improves mortality risk and reduces length of stay *Workforce:* Units – paediatric intensive care units (PICUs); tertiary care *Feature:* PICU volume – the number of admissions to the unit per year ranged from 147–1378. Volume scale for analysis is based on change of 100 admissions. *Outcome:* Mortality and length of stay
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Prospective cohort study 2 5 PICUs among 21 PICUs were excluded because they did not complete data collection. 3 11,106 consecutive patients in 16 PICUs 4 1 year; in units 5 Units' characteristics data were from Paediatric Critical Care Study Group (PCCSG); patients' data were from each unit's report for the study period; January 1993 to December 1993.
Results Quantitative results	An inverse relationship exists between patient volume and outcomes in the setting of the PICU. **Logistic regression results for PICU mortality** / **Negative binomial regression results for PICU length of stay** Volume: Regression coefficient –0.0005; SE 0.0002; *p*-value 0.045; Adjusted OR (95% CI) 0.95 (0.91–0.99) Volume: Regression coefficient –0.0002; SE 0.0001; *p*-value 0.030; Adjusted OR (95% CI) 0.980 (0.975–0.985)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for severity and case mix differences in mortality and length of stay by using Paediatric Risk of Mortality (PRISM) score and Paediatric Overall and Cerebral Performance scores as measures of functional status, and also adjusted for primary and secondary diagnosis codes, surgical and trauma status, and patient age. 2 Adjusted for university hospital affiliation, children's hospital, intermediate care unit, fellowship programme, and number of PICU beds. 3 Uniform 4 Complete 5 No 6 National
Commentary	This study did not choose study sites randomly based on pre-selected quality-of-care factors. The non-random selection of PICUs led to a greater number of high-volume academic institutions being included as study sites and may have created a sample selection effect. Data for this study were collected for an entire year, regardless of the number of deaths that occurred at any given PICU. Attempts to disentangle differences based on sample size calculations are confounded by adjustments for clustering. The study was unable to assess whether the estimated relationships provide evidence for or against either a practice-makes-perfect effect or a selective referral effect. The data do not permit an analysis of specific types of illness or injury cared for in PICUs.
Research implications	An investigation of specific types of illness or injury cared for in PICUs requires a broader set of clinical variables and a larger database. The low mortality in paediatric health services needs larger samples to estimate statistically significant relationships. A more comprehensive outcomes measure needs to include known complications and survival to hospital discharge. It is unclear whether smaller PICUs were less or more likely to use appropriate therapies. Providers can use this information in considering whether the system of care can be improved or whether patient care can be improved by referral to high-volume providers. Studies documenting improved outcomes from increased patient volume have led some researchers to speculate that volume will become the proxy indicator for quality in health care and will result in the reorganisation of systems to reflect this emphasis. A furthering understanding of this relationship is needed to develop effective regionalisation and referral policies for critically ill children.

ID, origin, authors (year)	114, UK, UK Neonatal Staffing Study Group (2002)
Aims	Whether patient volume, staffing levels, and workload are associated with risk-adjusted outcomes, and with costs or staff wellbeing. *Workforce:* Unit – neonatal intensive care units (NICU); tertiary care *Feature:* Unit patient volume – number of very low birthweight infants (<1500 g) admitted per year: <35, 35–57, >57 • Nursing provision: establishment nurse to cot ratios. <1 national minimum, 1–1.84 national minimum, >1.84 national minimum • Neonatal consultant provision: paediatricians, with more than 50% of their clinical session committed to neonatal care, per NICU <1, 1, ≥2 • NICU workload: percentage maximum occupancy and ratio of nurses to infants *Outcome:* Hospital mortality, mortality or cerebral damage, and nosocomial bacteraemia.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional observational study 2 Include units intending to provide sustained neonatal intensive care. Infants younger than 1 month were included. 3 13,334 infants consecutively admitted to 54 NICUs 4 One year; in-hospital 5 Data for infants at admission and discharge were taken from medical and nursing records by a trained link-research nurse at every NICU. Data for patient volume, nursing and consultant provision were obtained for every unit at 00.00 hours and 12.00 hours through the observation period, and were returned every week to the study centre. Data for resource use, including equipment, drugs, staffing, overheads, and recurrent cost data were gathered from all 54 units by an established questionnaire. March 1998 to April 1999.
Results Quantitative results	No association between patient volume and patient health outcomes. Mortality is directly related to initial occupancy and nurse-to-infant ratio in all types of units. Risk-adjusted nosocomial bacteraemia rose in NICUs with high consultant provision.

	Patient volume			**Nursing provision**		**Consultant provision**		**Occupancy**	**Nurse-to-infant ratio**
	high	***medium***	***low***	***high***	***low***	***high***	***low***		
Mortality									
Birth*	1.00**	1.12 (0.76–1.64)	0.97 (0.70–1.34)	1.00	1.00 (0.82–1.48)	1.00	0.92 (0.69–1.22)	1.09 (1.01–1.18)	0.98 (0.94–1.02)
12-hour	1.00	1.10 (0.75–1.62)	0.86 (0.60–1.23)	1.00	1.14 (0.85–1.53)	1.00	0.94 (0.71–1.25)	1.11 (1.02–1.20)	1.01 (0.96–1.06)
Mortality or brain damage									
Birth	1.00	1.10 (0.71–1.71)	0.99 (0.69–1.43)	1.00	0.97 (0.70–1.35)	1.00	0.96 (0.69–1.33)	1.03 (0.97–1.11)	0.99 (0.96–1.04)
12-hour	1.00	1.19 (0.77–1.83)	0.92 (0.65–1.30)	1.00	0.99 (0.71–1.37)	1.00	1.01 (0.73–1.38)	1.04 (0.97–1.11)	1.01 (0.97–1.06)
Nosocomial bacteraemia									
Birth	1.00	1.38 (0.85–2.22)	1.23 (0.70–2.15)	1.00	0.86 (0.56–1.32)	1.00	0.65 (0.43–.98)	0.99 (0.94–1.05)	0.99 (0.96–1.04)
12-hour	1.00	1.39 (0.87–2.21)	1.22 (0.70–2.12)	1.00	0.80 (0.53–1.20)	1.00	0.65 (0.44–.96)	1.00 (0.95–1.06)	1.00 (0.97–1.03)

* The study set up two care models: birth model – neonatal intensive care from birth; 12-hour model – neonatal intensive care up to 12 hours after admission.

** Data are odds ratios (95% CI)

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Birth model: For death in hospital, and death in hospital or cerebral abnormity, risk adjustment in the birth model consisted of the infant's gestation, size of infant for gestation, sex, mode of delivery, diagnostic category, clinical risk, initial illness severity and maternal treatment with antenatal steroids. For nosocomial bacteraemia, the adjustment also included whether the infant had an initial positive blood culture more than 48 hours after birth (excluding cases of probable vertical transmission). 12-hour model: additional adjustment for admission temperature, the most extreme $PaCO_2$, mean appropriate FiO_2, and the lowest base excess were also included. 2 No 3 Uniformed 4 Complete 5 No 6 National
Commentary	The confounding effect of an individual infant's own evolving nursing requirement might be possible. The ascertainment bias might have been caused by the variation in sampling frequency, variations in rates of contaminated blood cultures, or even that NICUs with more neonatal consultants could have been less successful in implementation of unified protocols for infection control. The omission on recording frequency of cranial ultrasound examinations and blood cultures could bias comparisons of mortality or cerebral damage and probable nosocomial bacteraemia in favour of units that did few investigations. The variations in observation periods will introduce the seasonal bias. The cohort of infants was not followed up beyond hospital outcome and no examination of differences in subsequent morbidity or psychomotor development was possible.
Research implications	Need to establish an optimum absolute value across the cohort for nurse-to-infant ratio in relation to risk-adjusted outcome. Future studies by daily measurement of a patient's nursing requirement could separate the individual infant's nursing requirement from the effect of total NICU nursing requirement. Need studies on measuring the required versus actual nurse provision per individual infant throughout stay in relation to outcome. Organisation of UK NICUs should aim to balance the conflicting demands of improving efficiency yet maintaining ease of access; it can be achieved by more formal development of co-operative neonatal networks, and reliable, agreed, and appropriate dependency categories with guidelines for transfer. Improvements to the service could be achieved by reduction of nursing workload. Important data on longer-term morbidity outcomes should be included in future assessments of neonatal intensive care.

ID, origin, authors (year)	1124, UK, University of York, NHS Centre for Reviews and Dissemination (1997)
Aims	To review the literature on the relationship between the volume of hospital or consultant activity and clinical outcomes. *Workforce*: Hospital; secondary care *Feature*: Hospital procedure volume *Outcomes*: Mortality (in-hospital or other), morbidity (e.g. infection rates), psychosocial (e.g. satisfaction), quality of life. (No studies assessing psychosocial outcomes were identified)
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 Inclusion: patients receiving any of the following: cardiovascular surgery (coronary artery bypass graft (CABG) surgery, other open heart, acute myocardial infarction and other heart problems, pacemaker implantation, cardiac catheterisation/angiography, percutaneous transluminal coronary angioplasty (PTCA), carotid endarterectomy (CE), abdominal aortic aneurysms, vascular and cerebro-vascular surgery); respiratory medicine; abdominal procedures (gastric operations, cholecystectomy, appendicectomy, intestinal hernia repair, gall bladder, ulcer); orthopaedic surgery (hip or knee arthroplasty, hip fracture); intensive care (neonatal/perinatal, paediatric, adult); urology/gynaecology (prostate, kidney/urinary tract infection and urology, hysterectomy, Caesarean section); trauma care; AIDS; cataract surgery; cancer; miscellaneous (for example patients with cirrhosis). 3 Included in 4 4 Not all studies presented data on number of hospitals or doctors. *CABG surgery:* 9 retrospective analyses (383,245 patients; 997 hospitals; 1,061 doctors) Open heart: 4 retrospective analyses (39,320 patients, 830 hospitals) *Myocardial infarction and other heart problems:* 1 prospective cohort (2,265 patients; 18 hospitals); 8 retrospective analyses (1,159,126 patients, 2,011 hospitals; 926 doctors) *Pacemaker implantation:* 2 retrospective analyses (201 patients; 1,752 hospitals) Cardiac catheterisation/angiography: 3 retrospective analyses (108,097 patients; 549 hospitals and 3,132 doctors); 1 survey (46,904 patients, 373 hospitals) *PTCA:* 1 RCT (50 patients); 5 retrospective analyses (267,591) patients; 1,457 hospitals; 38 doctors) CE: 1 pre-operation review (743 patients; 1 hospital; 24 doctors); 7 retrospective analyses (24,860 patients; 1,868 hospitals; 1,073 doctors) *Abdominal aortic aneurysms:* 1 prospective multi-centre cohort (444 patients, 26 hospitals); 9 retrospective analyses (39,825 patients; 2,305 hospitals; 874 doctors); 1 survey (294 patients, 17 hospitals) *Vascular and cerebro-vascular surgery:* 5 retrospective analyses (66,484 patients; 3,059 hospitals; 36 doctors) *Respiratory:* 4 retrospective analyses (10,425 patients) *Gastric operations:* 6 retrospective analyses (52,234 patients; 4,116 hospitals; 4,945 doctors) Cholecystectomy: 10 retrospective analyses (459,703 patients; 7,617 hospitals; 9,384 doctors) *Appendicectomy:* 4 retrospective analyses (132,122 patients; 1,676 hospitals; 6,434 doctors) Intestinal: 9 retrospective analyses (142,673 patients; 3,307 hospitals; 7,433 doctors) *Hernia repair:* 5 retrospective analyses (288,068 patients; 2,014 hospitals; 7,476 doctors) *Gall bladder:* 1 retrospective analysis (88,839 patients; 1,210 hospitals) *Ulcer:* 1 retrospective analysis (138,268 patients; 1,214 hospitals) *Hip or knee:* 11 retrospective analyses (237,508 patients; 4,384 hospitals; 2,700 doctors) *Hip fracture:* 5 retrospective analyses (146,233 patients; 4,534 hospitals) *Intensive care:* Neonatal/perinatal: 1 quasi experimental (matched control region) (7,394 patients); 1 before/after (87,213 patients); 2 prospective cohort (4,538 patients; 18 hospitals); 1 cohort (319 patients); 1 case control (1,179 patients; 39 hospitals); 1 case review (447 patients); 22 retrospective (6,038,834 patients; 2,219 hospitals; 715 doctors) Paediatric: 2 prospective cohort (5,878 patients; 90 hospitals) Adult: data from APACHE II study (11,612 patients, 26 hospitals) *Prostate:* 7 retrospective analyses (253,861 patients; 2,604 hospitals; 2,892 doctors)

	Kidney/urinary: 2 retrospective analyses (5,510 patients) *Hysterectomy:* 4 retrospective studies (297,740 patients; 1,673 hospitals, 8,027 doctors) *Caesarean section:* 1 retrospective study (3,478 patients; 22 hospitals) *Trauma care:* 1 prospective comparative (2,646 patients); 1 cross-sectional comparison (182 patients; 40 hospitals); 7 before/after (73,569 patients; 68 hospitals); 1 case–control (85 patients); 14 retrospective analyses (52,710 patients; 271 hospitals) *AIDS:* 2 retrospective analyses (557 patients; 55 hospitals) *Cataract surgery:* 1 stratified prospective cohort (772 patients; 75 doctors) *Cancer:* Breast: 2 retrospective analyses (17,873 patients; 180 doctors) Colorectal: 2 prospective cohort (1,239 patients; 7 hospitals; 56 doctors); 1 prospective audit (750 patients; 28 doctors); 3 retrospective analyses (23,781 patients; 1,156 hospitals; 434 doctors); 1 cohort (438 patients, 5 surgeons) Pancreatic: 1 RCT (145 patients; 1 hospital; 5 doctors); 1 retrospective analysis (1,972 patients; 184 hospitals; 748 doctors) Teratoma: 1 retrospective analysis (454 patients; 5 hospitals) Oesophageal: 1 retrospective analysis (1,143patients) Stomach: 1 retrospective analysis (341 patients; 69 hospitals; 193 doctors) Lung: 1 retrospective analysis (12,439 patients; 389 hospitals) Childhood: 2 retrospective analyses (4,438 patients) Oncologic procedures: 1 retrospective analysis (2,627 patients; 1 hospital) *Miscellaneous:* 5 retrospective (31,883 patients; 938 hospitals); 1 analysis of routine (3,434 patients; 14 hospitals). 5 MeSH Headings: MEDLINE (1980-1996), EMBASE (1974–1996), Health Planning and Administration (1975–1995), Dissertation Abstracts (1861–1996) and Entis. In addition key relevant journals were hand searched, references of identified studies were checked and experts in the field and other Health Technology Assessment bodies were contacted to help identify published and unpublished studies. 6 Type of study design, process of patient identification, degree of adjustment for patient case mix, avoidance of selection bias. The relevance of each individual study was assessed by one reviewer. The quality of each individual study was assessed by one reviewer. Patient case mix adjustment scores were allocated by two reviewers. Data were extracted in a systematic way by one reviewer. 7 A qualitative overview is presented, taking into account the methodological rigour of each individual study. Where studies are similar enough (e.g. procedure, volume measure, patient type and outcomes measured) formal pooling of the data has been attempted. Studies are grouped according to the procedure or condition and, within this, studies have been ranked according to the extent of adjustment for patient case mix. Differences were discussed in the narrative.

Results Quantitative results	Only the results of studies with adequate adjustment for case mix (Grade III) are presented here. OR = odds ratio (the ratio of the odds of an adverse event occurring in a higher volume unit compared to a low volume unit); if OR <1 then there is less risk of a poor outcome in the higher-volume unit. *CABG surgery:* Reduced risk of in-hospital mortality in hospitals carrying out >200 procedures/year (OR = 0.90). *Paediatric heart surgery:* Reduced death rate in hospitals with >300 case/year compared to hospitals with <10 cases and <300 cases (OR = 1/8 and 1/3 respectively). *Acute myocardial infarction:* No significant difference in in-hospital but higher 6-months mortality and lower rate of re-infarction in hospitals with <300 beds (mortality 17% vs. 12%). Significant negative relationship between in-hospital mortality and physician volume (coefficient = –0.05) but not hospital volume. *Cardiac catheterisation:* No physician volume relationship found. Mortality declines by 0.1% for a 100 increase in annual number of hospital procedures (average no. of treatments = 400). *PTCA:* No significant association between physician volume and angiographic or clinical success. Reduction in major complications when volume >400/year (OR = 0.66). No physician volume relationship found for mortality but more complications, emergency CABG and longer length of stay in physicians carrying out <50 procedures per year. 20% mortality for physicians <4 compared with 15% for physicians >4 procedures per year. *Abdominal aortic aneurysms:* SMR 30% higher in hospitals with <14 patients/year but no surgeon relationship found. 12% mortality for hospitals with <6 procedures compared to 5% in those >38 per year. Double the mortality in low-volume surgeons (<6) compared to high-volume surgeons (>26). Mortality declines by 1% for an increase of 4 operations per year per hospital (average no. of treatments = 23 per year). No evidence of a surgeon volume effect. 2% increased odds of dying if in hospital with <21 case compared to >21. This risk difference greater for ruptured aneurysms. *Amputation of lower limb (no trauma):* SMR 16% higher in hospitals with below-average annual volume (average no. of treatments = 10.5). *Gastric surgery:* No significant difference between hospitals with below- and above-average annual volume (average no. of treatments = 24). Mortality declines by 1% for a 17 increase in annual number of hospital operations (average no. of treatments = 38). No relationship between physician volume and mortality (average no. of treatments = 8). Surgeons carrying out <2 procedures annually associated with higher mortality rate than those doing >1. *Cholecystectomy:* SMR 26% higher in hospitals with below-average annual volume (average no. of treatments = 109). Hospitals performing <168 procedures a year had a mortality rate of 1.52% compared to 1.21% in those with higher volume. No significant association with surgeon volume found. *Intestinal operations (excluding cancer):* Hospital mortality higher (8.3%) when <40 operations per year than if >40 operations (5.9%). Surgeons with annual volume of >8 also associated with lower mortality. *Gall bladder (non-surgical):* SMR 14% lower in hospitals with below-average annual volume (average no. of treatments = 73). *Ulcer (non-surgical):* No statistically significant effect of volume. *Knee replacement:* Higher hospital volume associated with lower risk of complications (average no. of treatments = 3.5). *Hip fracture:* No significant effect of hospital volume on mortality (average no. of treatments = 45). *Neonatal care:* Infants <28 weeks gestation had better survival in intensive care units (>500 days ventilation/year) compared with special care units (<500 days of ventilation/year). No difference for more mature infants. *Paediatric intensive care:* No statistically significant association found between mortality and monthly volume. *Adult intensive care:* No association between % dying and monthly unit volume. *Prostatectomy:* No statistically significant differences found. *Trauma care:* No statistically significant association between mortality from major trauma and volume across A& E departments with volumes ranging from <10/year to >90/year in 3 regions with and without an experimental trauma system. No major differences in mortality in a tertiary trauma unit for patients with mainly blunt injuries as it doubled in volume over a 4-year period.

	Cataract surgery: Surgeons carrying out >200 operations a year had a greater rate of adverse events (especially posterior capsular opacification OR = 2.5) *AIDS:* Risk of 30-day mortality was 2.5 times as high when treated in low-experience hospitals (<43 patients) than in a hospital having treated >43 patients (RR for 30-day mortality = 2.5). *Breast cancer:* 15% reduction in mortality with surgeons treating >29 new cases/year but no advantage of >50 compared with >29. *Colorectal cancer:* SMR 20% higher in hospitals with below-average annual volume (average no. of treatments = 50) or surgeon volume (Average N of treatments = 8). *Laparotomy with colorectal resection:* No statistically significant differences in mortality or morbidity between surgeons with volumes ranging from 44 to 110 cases per year. *Stomach cancer:* No statistically significant association between mortality and either hospital or surgeon volume. *Malignant teratoma:* 5-year mortality 60% lower in patients treated in a cancer unit which treated over 50% of patients with this cancer in the area. *Oesophageal cancer:* 17% lower rate of operative mortality in surgeons performing <3 operations annually. 4% reduction in 5-year mortality with surgeons treating >5 new cases a year. Most explained by reduced operative deaths. *Pancreatic cancer:* Patients treated by surgeons with highest volume (76 cases in 20 months) had lowest risk of complications (fistula) compared to lower-volume surgeons in same hospital.
Commentary	The literature on links between volume of activity and clinical outcomes suggests that for some procedures or specialities there may be some quality gains as hospital or clinician volume increases. In other areas the research suggests an absence of significant volume gains. Generalisation is clearly not possible on the basis of these results. Hence it would not be warranted to extrapolate the findings; whether positive or negative, outside the sample ranges or for the many procedures where the research evidence is too poor to suggest any conclusion. Where volume is associated with quality, the direction of causation is not established and there is no good evidence to indicate that increasing volume will actually result in improvements in health care outcomes. This report was abstracted in conjunction with CRD Report 8 (1). This review was based on a clear research question. Thorough searches were carried. Both the inclusion criteria and methodological quality assessment are clearly defined in the full Report (8 part 1). This report combined the data in an appropriate manner and ranked in accordance with their case mix-adjustment scores. However, although case-mix scores were allocated by two reviewers, relevance and quality assessment and data- extraction were undertaken by only one. Double-checking is desirable as it reduces bias and errors.
Research implications	In the few cases where volume quality links have been suggested by more reliable studies, these might well act as prompts for investigation by purchasers and/or clinicians. In some cases, the indicated thresholds are relatively low and could be reached through specialisation of tasks within a hospital rather than through an increase in the size of the provider. There is a need for a well-designed case-study analysis of the effects of trust or hospital mergers (which have already taken place) on costs and clinical outcomes. There is a need for good-quality research to examine a broader range of indicators of outcome (such as quality of life or rates of re-admission or recurrence) and to establish the validity of the presumed benefits of sub-specialisation, multi-disciplinary working and inter-speciality links.

ID, origin, authors (year)	163, USA, Vakili, B.A., Kaplan, R. and Brown, D.L. (2001)
Aims	Whether the volume of primary angioplasty procedures for acute myocardial infarction (AMI) performed by physicians and/or hospitals is associated with a lower mortality rate. *Workforce:* Physicians and hospitals, angioplasty for AMI; secondary care *Feature:* Volume – number of angioplasty procedures performed per year. Physician volume 1–10, ≥11; hospital volume 1–56, ≥57 *Outcome:* In-hospital mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Include patients who underwent PTCA for AMI within 23 hours of symptom onset; exclude patients who had received thrombolytic therapy within 7 days before the procedure. 3 1342 patients undergoing elective and emergent PTCA in 32 participating hospitals 4 1 year; in-hospital 5 Data were from the Coronary Angioplasty Reporting System (CARS) of the New York State Department of Health (DOH); 1995
Results Quantitative results	An inverse relation exists between physician primary angioplasty experience and in-hospital mortality; a strong trend toward a relation between hospital primary angioplasty volume and mortality; an interaction between hospital and physician primary angioplasty volume and in-hospital mortality exists such that those AMI patients treated in high-volume hospitals by high-volume physicians have a 49% lower in-hospital mortality rate than those treated by low-volume physicians in low-volume hospitals. *Unadjusted and adjusted relative risk (95% CI) of in-hospital death among patients who underwent primary angioplasty for AMI according to physician and hospital volume* **High-volume vs. low-volume physicians** / **High-volume vs. low-volume hospitals** Unadjusted: 0.50 (0.3-0.77) / 0.67 (0.4-1.1) Adjusted for demographics: 0.53 (0.31-0.91) / 0.67 (0.4-1.1) Adjusted for demographics, medical history: 0.56 (0.32-0.91) / 0.67 (0.42-1.1) Adjusted for demographics, medical history, haemodynamic status, and time to treatment: 0.43 (0.21-0.83) / 0.56 (0.29-1.1) *Multivariate adjusted relation between physician annual primary angioplasty volume category, hospital annual primary angioplasty volume category, and in-hospital mortality (crude mortality rate; relative risk (95% CI))* **Low physician volume** / **High physician volume** Low hospital volume: 7.6% 1.0 / 5.8% 0.6 (0.21–1.73) High hospital volume: 4.1% 0.56 (0.24–1.28) / 3.7% 0.51 (0.26–0.99)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjusted for patients demographics (age and sex), and medical history (smoking, diabetes, previous MI, and previous cardiac surgery), haemodynamic status (patients requiring pharmacological support or the presence of systolic blood pressure <80 mmHg or a cardiac index <2 L/m2 despite pharmacological or mechanical support before commencement of procedure.) 2 Adjusted for patients' time to treatment (<6 hours or from 6 to 23 hours after symptom onset) 3 Uniform 4 Complete 5 No 6 One state

Commentary	The intervention requires more stable patients and more skilled practitioners, and the skilled practitioners are generally those with higher volume, which results in improved outcomes. The study's retrospective nature can only identify associations, rather than causality. Referral bias is possible when patients are attracted to doctors and hospitals because of their reputation for good results. Higher-risk patients are then disproportionately represented in lower-volume hospitals. The result of the studies might be affected by the changes in AMI treatment since 1995, which are not supported by RCTs. The CARS data set does not include information on adjunctive treatments that have been demonstrated to improve mortality from AMI. High volume hospitals are generally better at prescribing some medications.
Research implications	More studies needed to look at the relationship between physician volume and hospital volume. Different models for physician and hospital volume should be evaluated.

Table A2.11 Specialisation

ID, origin, authors (year)	172, USA, Alexander, F. *et al.* (2001)
Aims	To compare outcome of children with appendicitis cared by specialists versus generalist *Workforce:* Surgeons; academic medical centre *Feature:* Specialisation *Outcome:* Complications, re-admission, second operation and length of stay
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective, comparative study 2 Include all children 17 years of age or less who underwent emergency appendectomy at the Cleveland Clinic Children's Hospital and affiliate hospitals between March 1994 and December 1997. 3 175 children. Of those, 96 were treated by a Health Maintenance Organisation (HMO) Adult General Surgical Service (group A) and 79 were treated by a Paediatric Surgical Service (group B). 4 N/A 5 Treatment and outcome indicators were collected from hospital data. They included imaging tests performed, operation type, complications, re-admissions, and length of stay.
Results Quantitative results	Group A was HMO surgical staff comprising 6 general surgeons each with greater than 5 years of experience. Group B comprised 3 paediatric surgeons with a minimum of 2 years of experience. In patients with simple acute appendicitis, there was no significant outcome difference between groups A and B. In patients with gangrenous or perforated appendicitis, there were significant differences for complications (group A, 9 of 27 vs. group B, 3 of 34, p = 0.025); re-admissions (group A, 6 of 27 vs. group B, 0 of 34, p = 0.001); second operation (group A, 6 of 27 vs. group B, 2 of 34, p = 0.001); and mean total length of stay in days (group A, 8.6 of 27 vs. group B, 5.4 of 34, p = 0.05).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Group B had more patients with gangrenous or perforated appendicitis. Therefore, patients were categorised into two groups: simple acute appendicitis and gangrenous/perforated appendicitis in order to give a better comparison. 2 No 3 Yes 4 Completed 5 No. The assignment of patients to group A or group B physicians depended on whether the patients were enrolled in that HMO. Those who were in HMO were treated by group A. All other children, including those insured by Medicaid or third-party payer, were treated by paediatric surgical staff (group B). 6 In Cleveland Clinic Children's Hospital and affiliate hospitals in Cleveland, OH
Commentary	Small sample size. It will be better if the study includes more participating hospitals.
Research implications	Future research may include cost-effectiveness in addition to outcome measures.

ID, origin, authors (year)	11, USA, Anderson, J.J. *et al.* (2002)
Aims	To evaluate costs and effectiveness of ambulatory care provided by specialists, non-specialists (general internists) and both specialists and non-specialists (co-care) to outpatients with knee osteoarthritis (OA) and/or chronic low back pain (LBP) *Workforce:* Specialists and internists; Veteran Health Administration *Feature:* Specialisation *Outcome:* Physical Component Summary (PCS, to measure functional status) and costs
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective, comparative study 2 Included patients who had a baseline interview between August 1993 and June 1995, at least one subsequent OA- or LBP-related visit to a Veterans Health Administration (VHA) internist, rheumatologist, orthopaedic surgeon or neurologist, and a subsequent quarterly medical outcome study short form 36-item questionnaires as part of the Veterans Health Study (VHS) with follow-up time of at least 6 months between the baseline interview and December 1995. 3 398 patients (155 received only non-specialty care, 49 specialty-only care and 192 co-care) 4 An average of 14 months 5 Obtained VHA outpatient utilisation data for each patient from the Decentralised Hospital Computer Program (DHCP) which include information on all patient visits, laboratory tests and pharmacy between the baseline interview and December 1995 so that utilisation costs could be estimated; patients completed medical outcomes study short form 36-item functional status questionnaires at both baseline and follow-up.
Results Quantitative results	The effectiveness of ambulatory care provided by specialists and non-specialists was examined. Physical Component Summary (PCS) improvements per year were 1.66 (SD 8.22) for non-specialty care, 3.48 (SD 7.91) for specialty care, and 0.65 (SD 8.08) for co-care while costs of care per year were $1099, $1376 and $2517, respectively. A standardised incremental cost-effectiveness ratio (ICER) of $152 per PCS unit indicated specialty care to be cost-effective compared with non-specialty care.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 PCS improvement score was adjusted for age, disease characteristics, and baseline health status. 2 The stability of each ICER was assessed by bootstrap sampling, a repeated sampling technique used to provide a nonparametric estimate of the distribution of an ICER to reduce susceptibility to effects of possible outliers. 3 Yes 4 Completed. Patients were followed up for at least 6 months (an average of 14 months). 5 Patients were not randomly assigned to be cared by a specialty or non-specialty. However, in VHS, participants were random sample of patients who came to one of 4 VHA sites for an ambulatory care medical visit between August 1993 and December 95. Therefore, participants in this study obtained care from a wide range of different individual providers in the Boston-area VHA centres. 6 Not dispersed. 4 sites in Boston.
Commentary	By design, this study is limited to the VHA outpatient setting. Not all care provided to veterans for OA or LBP was included. Inpatient care, either within or outside the VHA, was not included, and some veterans may also have received outpatient care for the conditions outside the VHA.
Research implications	In VHA outpatient care in 4 Boston area clinics in 1994–1995, specialist-only care resulted in improved functional status outcomes for patients with OA or LBP.

ID, origin, authors (year)	1140, USA, Ayanian, J.Z., Landrum, M.B., Guadagnoli, E. and Gaccione, P. (2002)
Aims	To evaluate the relation between ambulatory care and mortality among elderly patients after myocardial infarction. *Workforce:* Ambulatory care physicians: cardiologists, internists, and family practitioners *Feature:* Training of workforce *Outcome:* Patient mortality (measured by a 2-year morality rate). The study also evaluated the number of office visits patients had after the MI, and other characteristics of patients such as sex, race, conditions before admission, clinical complications in hospital, type of hospital and type of care; however, these were not connected to patient mortality after ambulatory care so they will not be reported in this abstraction.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional 2 52,064 patients 65–84 years of age with fee-for-service Medicare who were discharged alive after a clinically confirmed MI. Excluded patients who died with in three months after discharge, those who had metastatic cancer or a DNR, those enrolled in an HMO within three months after discharge, and those who resided in nursing homes or who lacked Medicare Part B coverage for physician's care. From the remaining 42,971 patients they excluded patients who did not report visiting an ambulatory care physician within 3 months of discharge and those whose clinical data were incomplete. 3 35,520 patients from 7 states (CA, TX, FL, MA, NY, OH, PA) 4 Data from Cooperative Cardiovascular Project, Medicare Part A, hospital records and hospital outpatient claims were used to identify elderly Medicare patients who were hospitalised with a principal diagnosis of MI during 1994–1995 who sought some type of coronary procedure 3 months after discharge as well as their demographics, coexisting illnesses, cardiac complications, test results and cardiovascular medications. The use of cardiovascular medications and other symptom controls approx. 18 months after discharge was assessed by a telephone survey. Hospital characteristics were gathered from Medicare and AHA data. Ambulatory visits to physicians were determined from Medicare Part B and hospital outpatient claims. Paid claims were identified with Current Procedural Terminology (CPT-4).
Results Quantitative results	Ambulatory care by cardiologists was associated with a lower mortality among elderly patients, and a further reduction in mortality was noted among patients treated by both cardiologists and internists or family practitioners. The 2-year mortality for the unmatched cohort was 11.8% for those who saw a cardiologist in the first 3 months after discharge and 19.1% for those who saw only an internist or a family practitioner (P<0.001). This absolute difference in mortality of 7.3% was reduced by half, to 3.7% (14.6% vs. 18.3%) after matching but remained statistically significant (p <0.001). The absolute reduction in mortality associated with cardiology care was greatest among patients with the least propensity to visit a cardiologist. Among patients in the unmatched cohort who visited both a cardiologist and an internist or family practitioner, the 2-year mortality rate was not significantly lower for those who just visited a cardiologist (11.5% vs. 12.2%, p =0.12). However, after matching the difference in mortality rates was significant (11.1% vs. 12.1%, p =0.02). Sensitivity analysis estimated the effect of controlling for an unmeasured variable such as high school degree that would have increased the likelihood of visiting a cardiologist by 10%, and could have been associated with a 40% reduction in mortality. Adjusting for such a factor would reduce the absolute difference in mortality between patients who did and did not see a cardiologist from 3.7 to 2.8%, but it still would remain statistically significant. However, if an unobserved variable were associated with a 10% relative increase in the rate of concurrent care by cardiologist and generalist physicians, the absolute difference in 2-year mortality between those who only saw a cardiologist and those who had concurrent care would be insignificant.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patients were matched for the likelihood that they would receive ambulatory cardiology care using a propensity-score methods as a function of 36 variables, including patients' demographic and clinical characteristics, care provided in hospitals, medications at discharge, and hospital characteristics. 2 Other adjustments were estimated on the mortality rate for controlling unmeasured factors such as high school degree that could have an effect on mortality rates. 3 Yes 4 Yes, up to two years after MI 5 No 6 This study represents fee-for-service Medicare patients who are between the ages of 65 and 84 who suffered MI in the states of California, Florida, Texas, Ohio, New York, Pennsylvania, and Massachusetts
Commentary	Relied heavily on Medicare data; data on the use of cardiovascular drugs were available for only a sample of patients who completed the telephone survey; no data on coronary procedures performed more than 3 months after discharge; excluded patients who were enrolled in HMOs where the effects of primary and secondary care may be different from those with fee-for-service. The study provides insight into the characteristics of people who seek ambulatory care from cardiologists and/or from generalists. The study used a large representative cohort; a longitudinal assessment of Medicare claims for ambulatory care, and rigorous propensity-score methods to minimise selection bias.
Research implications	What is the relationship between ambulatory care and mortality among elderly patients after myocardial infarction who are enrolled in HMOs? What is the relationship between number of office visits after MI and patient mortality? Is there a relationship between ambulatory care and symptom control among elderly patients after MI? Why do patients who visit cardiologists have lower mortality rates – what are the specifics?

ID, origin, authors (year)	82, Portugal, Azevedo, A. (2002)
Aims	To assess the effect of outpatient management at a heart failure (HF) clinic, as compared with care by the usual assistant physician, on prognosis of HF patients. *Workforce:* Staff of HF-specific outpatient clinic and assistive physician (personal care physician) *Feature:* Use of therapeutic guidelines, drugs/agents in care of HF patients *Outcome:* Comparison of prognosis to all causes of death, cardiac-cause re-hospitalisation, long-term survival
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective non-randomised cohort study 2 All patients with HF defined according to guidelines for diagnosis, discharged from medical ward of community hospital. No data on exclusions. 3 339 patients 4 1-month, 6-month follow-up by mail and phone to 700 days 5 Source of data was medical records collected from one hospital for the 2-year period between January 1995 and December 1996.
Results Quantitative results	The risk of dying or being readmitted during the first month after discharge was significantly lower in patients followed at the HF clinic (adjusted odds ratio 0.23; 95% CI 0.12–0.46). Patients followed in the HF clinic also had an independent significantly lower hazard of dying during a longer-term follow-up of average length 373 days (adjusted hazard ratio 0.52; 95% CI 0.34–0.81). The results support the fact that a multidisciplinary and permanently available medical staff might be of relevance in improving outcomes in HF patients.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment 2 Adjusted odds ratio 3 Uniform data collection 4 Non-random 5 One clinic and one community hospital setting
Commentary	The HF outpatient clinic allowed patients regular appointments and unscheduled visits or phone consultations when needed. There were internists and cardiologist, a trained nutritionist and specialty nursing staff. Prior studies (e.g. the Diabetes Control and Complications Trial) have shown intensity of care for chronic patients results in better long-term outcomes and patient compliance to regimes.
Research implications	Relevance of multidisciplinary and permanently available medical staff in improving outcomes in HF patients.

ID, origin, authors (year)	178, Italy, Bellelli, G. *et al.* (2001)
Aims	To determine whether the use of different kinds of physicians leads to different outcomes with regard to the rate of hospitalisation and appropriateness of the management of adverse clinical events (ACE) in Italian nursing homes at night and during holidays. *Workforce:* Staff physicians (SP), temporary physicians (TP), and publicly funded National Health System (NHS) physicians *Feature:* Medical intervention during ACEs which occur during night and holiday periods *Outcome:* Hospitalisation rate, appropriateness of management
Methods	
1 Design	1 Prospective, non randomised-survey data collection
2 In-/exclusion	2 Geriatric populations at 10 non-profit nursing facilities in Lombardia Italy (431 nursing homes in area)
3 Sample size	3 352 nursing home residents had 551 adverse clinical events: 78 patients were hospitalised.
4 Follow-up time	4 1 day post-ACE follow-up
5 Data collection: source and period	5 Data were collected from data forms filled out by physicians at the time of the ACE, from NH administrative charts and medical charts from 14 months of holiday and night (April 1996–June 1997).
Results Quantitative results	The hospitalisation rate of NHS physicians was about twice that of the temporary physicians and six times that of the staff physicians. Probability of hospitalisation as percentage (95% CI): SP = 8.7% (2.3–15.1), TP = 11.2% (5.2–7.2) and NHS 32.2% (24.0–40.1). Staff physician' diagnoses and management were appropriate in the majority of cases, NHS diagnosis and management were doubtful or incorrect in about one-third of all cases. Therefore, nursing home residents frequently experience adverse clinical events; physician characteristics influence the rate of hospitalisation and quality of medical interventions.
Quality appraisal	
1 Case mix adjustment	1 No case-mix adjustment
2 Other adjustment	2 No other adjustment
3 Uniform data collection	3 Uniform data collection
4 Participant follow-up	4 1 day follow-up charting
5 Random sampling	5 No randomisation
6 Geographical dispersal	6 10 non-profit nursing home facilities in Lombardia, Italy, of a total of 431 in area
Commentary	Limitations: No initial assessment of basic characteristics of NH residents and whether different nursing home had a more frail population. The study demonstrates the favourable effects on hospital admission rates when physicians providing urgent care to NH residents are members of the internal staff rather than TPs or NHS physicians. A more intensive level of care should be encouraged within nursing homes.
Research implications	More research to evaluate optimal quality of level of care for residents of nursing homes on off-cycle hours.

ID, origin, authors (year)	607, UK, Bellingan, G., Olivier, T., Batson, S. and Webb, A. (2000)
Aims	To evaluate the effect of transfer method on acute physiology and early mortality. *Workforce:* Specialist doctors, junior doctors and nurses, and allied (ICU transport), primary care *Feature:* Training of workforce *Outcome:* Health status measured by acute physiology (pH, PaO_2, $PaCO_2$, heart rate (HR), mean arterial blood pressure (MAP)), and APACHE II/ SAPS II scores (with in 2 hours of admission,) and early mortality (<12 hours after admission)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Retrospective observational 2 Looked at all patients who were transferred into the University College London Hospitals (UCLH) intensive care unit from 1 October 1996 to 30 September 1997 3 259 patients from UCLH, either general, specialist, or teaching hospitals 4 In-hospital 5 Data is assumed to be collected from UCLH records
Results Quantitative results	Researchers found that the specialist teams were less acidotic and hypertensive upon arrival and, had a lower mortality than the standard ambulance team. There were no differences in the overall severity of illness (APACHE II and SAPS II scores). Group A* had a mean of 17.2 ± 7.4 and 31.7 ± 13.6 respectively. Group B** had a mean of 17.8 ± 8.0 and 33.7 ± 17.1 respectively. There was no difference in acute physiology scores except for pH and MAP scores (p <0.05). pH <7.1: n (%) p-value group A/ group B: 5 (3.0)/10 (11.0) p = 0.008 MAP <60mm Hg: n (%) p-value group A/ group B: 15 (8.9)/16 (17.6) p = 0.03 Group B had more deaths within 6 hours of admission – 4 deaths, including one en route, of the 91 transferred (4.4%) than group A with only one death amongst the 168 transferred (0.6%). This difference was maintained up to 12 hours after admission, with 7.7% or group B patients dying compared with only 3% of those in group A. * Group A: Mobile ICU consisting of a trained doctor, nurse, driver, and medical physics tech. All trained in the transfer of ICU patients. ** Group B: Standard emergency ambulance team with medical escort provided by the referring hospita
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patients were controlled for demographics and acute physiology but no adjustments were made mainly because there were no differences. 2 Controlled for hospital type, time in referring hospital prior to transfer, admission diagnosis and ICU mortality per diagnosis but no differences were found between the two groups. 3 Yes 4 In-hospital 5 No. 6 This study represents all UCLH.

Commentary	Too few deaths to report the Mantel-Cox log rank test showing a statistically significant difference at 6 hours (p = 0.03, with p = 0.07 at 12 hours) for fewer early deaths in group A. Group A included more medical patients and more patients transferred from other ICU, both known to have worse prognosis. Although group A faired better, the results would be more accurate if they had controlled for this. Although stated as an outcome measure, length of stay was not reported in the results. Difference in pH and MAP reflect the degree to which patients are resuscitated, which is influenced by both by the sophistication of available monitoring and by the experience of staff interpreting these data. The fact that no differences were seen between groups in the oxygenation suggests that pulse oximetry monitoring, which is almost universally employed and easily interpreted, is of positive benefit and made these results more valid.
Research implications	It would be beneficial to make this study a cross-sectional study to represent other hospitals in England and/or UK; this will also make the sample size much larger. Study should be repeated adjusting for numbers of medical patients and patients transferred from other ICUs. Does transfer method have an effect on length of stay?

ID, origin, authors (year)	149, USA, Bini, E.J. *et al.* (2001)
Aims	To compare the length of hospital stay, cost of hospitalisation, and outcomes when generalists work together with gastroenterologists or alone in the management of patients admitted to the hospital with decompensated cirrhosis. *Workforce:* GI consultant and personal care physician, vs. personal care physician only *Feature:* Length of time to consultation (to 72 hours), management of patients *Outcome:* LOS, cost re-admission rate and mortality rate
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective cohort study 2 Inclusion: consecutive patients admitted to a single Veterans Affairs (VA) hospital with decompensated cirrhosis as part of a larger study of 2,320 admissions Exclusion: cirrhotic patients who were admitted for other reasons, such as pneumonia or chest pain, also patients admitted for acute gastrointestinal bleeding 3 N=197 patients; of these, 107 patients had a GI consult. 4 30–day, and long term to 21 December 2000. Median follow-up time 618 days. 5 Data were admitting office data and medical service admission logs collected over a one-year period between April 1998 and 6 March 1999. Long-term follow-up continued to 31 December 2000.
Results Quantitative results	Patients who had a GI consultation had a significantly shorter length of stay (5.6± 3.5 vs. 10.1± 5.8 days, p <0.001) and a lower cost of hospitalisation ($6,004 ± $4,994 vs. $10,006 ± $6,183, p <0.001) than those patients managed by generalists alone. The 30-day incidence of re-admission (13.3% vs. 27.8%, p = 0.01) and mortality (7.5% vs. 16.7%, p = 0.045) were significantly lower in the GI consultation group. During the median follow-up period of 618 days (range, 2–970), patients who had a GI consultation during hospitalisation had a significantly longer time to hospital re-admission (p <0.001) and improved survival (p = 0.02).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Case mix adjustment for age, gender and comorbid disease 2 Adjustment for aetiology of liver disease, time of year, admitting diagnosis, individual ward attendings' specialty 3 Uniform data collection from three sources: standardised data collection sheets at admission, hospital electronic records and GI consultation logs. 4 Yes, to median of 618 days 5 Non-randomised 6 1 VA teaching hospital site affiliated with New York University School of Medicine
Commentary	The GI division of the hospital is strictly consulting; patients are treated by a team of interns and residents under supervision of a physician. Many of the GI patients were more likely to have been seen in a GI clinic in the year prior. This and the higher severity of illness in this portion of the cohort may have influenced the request for a GI consultation. Limitations: Only one hospital site, and VA system is not fee for service.
Research implications	There is a need for additional studies to evaluate the impact of GI consultation on the outcome of patients with chronic liver disease, particularly to evaluate whether GI consultation was associated with improved health-related quality of life.

ID, origin, authors (year)	210, UK/Norway, Bradley, P. and Lindsay, B. (2002)
Aims	To compare the effectiveness of specialist epilepsy nurses in improving patient care with routine care *Workforce:* Nurses *Feature:* Specialisation *Outcome:* Patient care
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 *Inclusion criteria:* Study type: Randomised controlled trials and quasi-randomised controlled trials. Subjects: Of any age or sex, referred with a suspected new diagnosis of epilepsy or with an established diagnosis of epilepsy. Intervention types: (i) Specialist epilepsy nurse care, from nurses trained to manage the problems of people with epilepsy; (ii) routine care, defined as care received in general practice or hospital, which does not involve the services of a specialist epilepsy nurse. Outcomes: Those linked to patient's quality of life following nurse intervention, with or without the use of proxy measures. Suitable outcomes for our scoping study include: seizure frequency; appropriateness of medication prescribed; social or psychological functioning scores; objective measures of general health status or quality of life; number of days spent on sick leave or missing school; adverse effects. Other outcomes not applicable to our scoping study: Knowledge about epilepsy scores; employment status; costs of care; patients'reports of information received 3 Four trial reports relating to three trials 4 RCTs in general practice setting (1 study); RCTs in hospital setting (2 studies) 5 MEDLINE (Ovid): 1966–2002; Cochrane Controlled Trials Register (The Cochrane Library Issue 3, 2002); EMBASE 1988 to August 2002; PsycINFO (WebSPIRS 5), 1996 to September 2002; CINAHL (Sliverplatter) 1982 to June 2002; HealthSTAR: 1992–1999 Via NLM. Also searched: GEARS; ECRI; Effectiveness Healthcare Bulletin; Effectiveness Matters; Bandolier; Evidence Based Purchasing; National Research Register for ongoing research; vignettes. Additional sources of information: expert panels from Standing Group on Health Technology Assessment; experts in the field; references on papers already received; web sites. 6 Quality assessment was based on the adequacy of allocation concealment. Studies were assigned to: concealment adequate (allocation by telephone randomisation etc.); concealment unclear – necessitating contact with authors; inadequate concealment (allocation by day of week etc.). 7 No formal pooling or meta-analysis of studies attempted due to sufficient clinical heterogeneity found on reviewing differences across the trials.

Results Quantitative results	Results were described as outcomes given elsewhere in this abstract, and are presented specifically for the three included studies referenced by the authors: • Seizure frequency (1 study): No difference found in the number of people having no seizures or one or fewer seizures per month between groups as measured in the first 6 months after intervention (p = 0.494). • Appropriateness of medication prescribed (1 study): Epilepsy nurse found 11.1% of participants required medication management changes, although no information given on whether these proposed changes were or were not appropriate, nor is their any control group comparison. • Psychosocial functioning scores, depression and anxiety (2 studies): The HAD scale was used to assess this outcome as administered by hospital-based nurses. One study found no overall difference in anxiety (p = 0.635) and depression (p = 0.500) between control and intervention groups at 6 months. Another study found no significant difference between control and intervention groups in either anxiety (p = 0.41) or depression (p = 0.27), but there was a trend towards improvement. • Social functioning (1 study): The Impact of Epilepsy scale was used to assess this outcome and there was no siginificant difference found between control and intervention groups in social outcomes at six months (p = 0.125 after adjustment for sex and employment status) • Health status scores at end of follow-up (1 study): The EuroQol was used to assess this outcome. No difference between study and intervention groups with respect to overall health status, as measured by weighted health status (p = 0.496) or self-related health status (p = 0.364). • Sick leave, school absence at end of follow-up (1 study): No difference was found in the number of days' absence from work in the control and intervention groups (p = 0.864) at 6 months. • Adverse effects were not reported in any of the trials.
Commentary	Sufficient details of included and excluded studies provided. Primary studies were summarised appropriately – descriptively with p-values for quantitative outcomes; however, the review loses impact due to the heterogeneity of the studies. Two reviewers independently assessed the studies for inclusion, resolving disagreements in conference and the same two reviewers extracted the data.
Research implications	Review states that there is a paucity of research on the effectiveness of specialist nurses. Present studies are small in number and consider heterogenous populations. There are very few studies of high quality. Studies involving several specialist nurses are needed, and research should continue to identify subgroups of epilepsy sufferers who might benefit most from interventions.

ID, origin, authors (year)	137, UK, Campbell, W.B. *et al.* (2001)
Aims	To document the medical status of amputees clearly, and to demonstrate the effect of co-morbidity or mortality and examine the effect of surgeon seniority on outcome. *Workforce:* Consultants, registrars, senior house officers performing amputations *Feature:* American Society of Anaesthesiologists (ASA) grades, prior attempts at revascularisation, and seniority of surgeon *Outcome:* Amputation level, revision, complications and death
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Historical cohort review 2 Inclusion: Case notes of 349 consecutive primary (major lower limb) amputations in 312 (163 male) patients. Exclusion: 22 patients without case notes. 3 349 surgical amputations 4 30-day mortality follow-up 5 Computerised database of proformas completed by hand after surgery from medical records that comprised discharge documents collected over the period from 1992 to 1998
Results Quantitative results	The majority of patients were ASA 3 or 4 (76%) and ASA 4 was associated with increased mortality (p <01). Limiting heart problems (p <.01) and 'general frailty' (p <.001) also carried significantly higher risks, but limiting chest problems, dementia, and diabetes mellitus did not. There was no significant association between attempts at revascularisation at any time before amputation, and amputation level or the need for revision. There were no differences between consultants, registrars, and, senior house officers (most senior surgeon) for any outcome measure.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Adjustment for differences between groups of patients by chi-squared testing. Comorbidities and mortality recorded for each operation. 2 'Local' complications and 'remote/general' complications analysed separately. 3 Yes, proformas completed by study authors from medical records. One method of data collection. 4 N/A 5 Non-random 6 Not documented.
Commentary	It is unclear whether the results are from a single hospital or a wider set of sites. Retrospective case note review may not have detected all major complications.
Research implications	There is a trend towards increased complications for patients with cardiac and chest disease, ASA grade 4, and 'frail' patients. Further study of attempted revascularisation and its impact on final amputation outcome would be useful.

ID, origin, authors (year)	697, USA, Casale, P.N. *et al.* (1998)
Aims	To determine the effect of specialty care on in-hospital mortality in patients with acute myocardial infarction (AMI) *Workforce:* Physicians and cardiologists; secondary *Feature:* Specialisation *Outcome:* In-hospital mortality and length of stay
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Comparative retrospective study 2 Included all direct hospital admissions for the treatment of AMI in Pennsylvania in 1993 except 510 patients who were excluded from the analysis, who were under the age of 30 or over the age of 99, patients who left against medical advice, patients of clinical complexity, patients involved in hospital transfers, patients of physicians who treated more than 100 patients in which it appeared that the entire group's cases were assigned to a single physician, patients treated by a specialty other than cardiology or primary care and patients treated at a hospital that closed since 1993 or at a hospital that treated fewer than 30 AMI patients in 1993. 3 30,715 admissions – 510 excluded = 30,205 admissions 4 N/A 5 Hospital admissions in 1993 were identified from the International Classification of Disease (ICD-9-CM) code for myocardial infarction, initial episode of care as the principal diagnosis. Then all acute care hospitals were required to abstract previously established key clinical data from all patient admissions. The assignment of specialty was self-reported by the individual attending physician.
Results Quantitative results	The effect of specialty care on in-hospital mortality in AMI patients was studied. After adjustment for patient characteristics, a multiple logistic regression analysis identified treatment by a cardiologist (odds ratio = 0.83, p <0.003) and physicians treating a high volume of acute myocardial infarction patients (odds ratio = 0.89, p <0.03) as independent predictors of lower in-hospital mortality. Treatment by a cardiologist as compared to primary care physician was also associated with a significantly lower length of stay for medically treated patients (p <0.01). (Treatment by a cardiologist is associated with approximately a 17% reduction in hospital mortality in AMI patients.)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 A risk-adjusted model of in-hospital mortality was developed for the patients admitted directly to a hospital by testing 20 clinical and demographic variables, including the Atlas admission severity score which itself is a collection of 23 clinical variables. 2 510 patients were excluded form the analysis (see Methods 2). 3 Individual hospitals may have slight differences in recording patient's information. 4 Completed. 5 Include all AMI admissions in 1993 except 510 patients. The assignment of attending physicians (whether cardiologist or primary care) was the decision of the hospital and its physicians. 6 All admissions in one state (Pennsylvania)
Commentary	The study samples were from only one state but the sample size was large.
Research implications	Cardiologists have lower patient mortality and length of stay compared to primary care physicians. These results have important implications of optimal treatment of AMI in the current transformation of the health care delivery system.

ID, origin, authors (year)	251, USA, Chen, J. *et al.* (2000)
Aims	To examine whether better survival rates for patients suffering acute myocardial infarction (MI) attributed to specialty care can be attributed to other patient characteristics, e.g. comorbidity and functional limitations. *Workforce:* US physicians of cardiology patients: comparison between cardiologists, internal medicine subspecialists, family practitioners and general practitioners *Feature:* Physician specialisation, use of guideline-supported therapies, differences in clinical characteristics of patients *Outcome:* Mortality: in-hospital, 30 days and 1 year
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Comparative retrospective cohort study 2 Inclusion: Medicare patients hospitalised for MI between the years 1994 and 1995. Exclusions: patients for whom MI was not clinically confirmed; patients of age <65 years; patients transferred from another hospital; patients with terminal illness; patients for whom vital status at year 1 was not known, patients for whom this was not the first hospitalisation for MI; patients hospitalised outside the USA; patients whose records or physician records could not be merged with AHA Hospital Database or AMA Physician Masterfile; patients who were not treated by a physician who was not board-certified physician or whose physicians were of specialties other than cardiology, internal medicine, other internists, or family or general practice. Patient files from Alabama, Iowa, Wisconsin and Minnesota were excluded as they did not contain demographic data and had limited clinical information on all hospitalisations. 3 109,243 Medicare patients hospitalised for MI: 37,876 treated by cardiologist; 13,693 treated by sub-specialist in internal medicine; 31,809 treated by internist; 21,016 treated by family practitioner; 4,849 treated by a generalist. 4 N/A 5 Data from Medicare patients hospitalised between 1994 and 1995 from the national Cooperative Cardiovascular Project. Study was a survey of patient records. Patient records were reviewed for diagnosis, demographics, clinical variables, comorbid conditions, functional limitations, and physician specialty.
Results Quantitative results	Patients who had board-certified cardiologists as attending physicians had the least number of comorbid conditions, whereas patients who had general practitioners or internal medicine subspecialists as attending physicians usually had the most comorbidities. Cardiologists had the greatest use of most guideline-supported therapies, and general practitioners had the lowest use. After adjustment for severity of myocardial infarction, clinical presentation, and hospital characteristics, patients treated by cardiologists were less likely to die within 1 year (RR = 0.92, 95% CI: 0.89–0.95), and patients cared for by other general practitioners were more likely to die within 1 year (RR = 1.09, 95% CI: 1.03 to 1.14), than patients cared for by general internists. After adjusting for additional measures of comorbid illness and functional limitations, the 1-year survival benefit associated with cardiology care was attenuated relative to internists (RR = 0.97, 95% CI: 0.94 to 1.0) and the excess mortality associated with general practitioners decreased (RR = 1.05, 95% CI: 1.00 to 1.11). After adjustment for utilisation of guideline-supported therapies, differences in 1-year survival between patients treated by cardiologists, general practitioners and internists were not statistically significant.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjustment for severity of MI, clinical presentation and hospital characteristics. Chi-square tests for categorical variables and analysis of variance for continuous. 2 Adjustment for additional measures of comorbid illness, functional limitations and use of guideline-supported therapies 3 One method of data collection: within-method uniformity 4 100% – all selected patient files included to 30-day and 1-year points. 5 Not random. Types of physicians dependent on hospital characteristics. 6 National study, with some exclusion of central states without comprehensive patient data. Multiple hospital sites, but no sites specified.

Commentary	This study was designed specifically to determine whether patients with acute MI who were treated by cardiologists have better outcomes than patients treated by generalists. The authors adjusted for baseline differences in patient characteristics, case-base mix of comorbid illness and functional limitations, hospital and physician type. It seems clear that cardiac specialists get the 'ideal' MI candidates and that the extent of their knowledge and use of current guideline-supported therapies may account for better patient outcomes.
Research implications	The authors report equivalent outcomes for specialist and generalist physician populations once patient comorbidities and functional limitations are factored in.

ID, origin, authors (year)	398, USA, Czaplinski, C. and Diers, D. (1998)
Aims	To examine the effect of concentrated staff nurse expertise (a nursing specialty unit) on patient outcomes of length of stay and mortality. *Workforce:* Staff nurses working in specialised units *Feature:* Accumulated knowledge/specialisation through accumulated knowledge *Outcome:* LOS, mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Inclusion: All patients from the period classified by 16 Diagnosis Related Groups (DRG). Exclusion: DRG outliers were 'trimmed', and ICUs were removed from the sample. 3 N=11,316 discharges, 13 specialty units and 518 physicians 4 In-hospital data only, no follow-up 5 Clinical patient data were collected from the computerised records of one 800-bed teaching hospital over the 7-year period from 1987 to 1993.
Results Quantitative results	Qualitative results: Staff nurse specialisation decreased length of stay in 13 DRGs (after removing outliers). Mortality on specialised units was lower
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Each DRG treated as sub sample, adjustment for age. 2 No adjustment for physician specialty 3 Not specified. 4 N/A 5 No 6 No, single teaching hospital that has had specialty units for more than 30 years
Commentary	For the purpose of measuring outcomes (clinical or financial) it is possible that the specialised units may have sicker patients admitted. The larger the hospital unit the more likely it will be to have a nonspecialised case mix because it will get overflow from other services. Understanding LOS and mortality is becoming increasingly dependent on more intricate analysis of how the institution works: the relation among physician specialty, case mix by hospital unit, and nursing care.
Research implications	Specialised expertise as acquired on a specialised hospital unit is the equivalent of the standards embedded in clinical pathways in use today. It would be worth tracking quality indicators such as incident reports, length of time to extubation, variances from critical paths and other contemporary measures of process to examine the relationship between structure (nursing care delivery) and outcome.

ID, origin, authors (year)	779, UK, Dale, J. *et al.* (1996)
Aims	To compare outcome and costs of general practitioners, senior house officers, and registrars treating patients who attended accident and emergency department with problems assessed at triage as being of primary care type. *Workforce:* General practitioners, senior house officers and registrars *Feature:* Specialisation *Outcome:* Patient satisfaction (satisfaction withclinical assessment, treatment and consulting doctor's manner) and status (fully recovered, improving)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective intervention study 2 Included all patients presenting with primary care problems in 419 3-hours sessions throughout a 12-month period. 3 4641 patients (1702 were seen by general practitioners, 2382 by senior house officers, 557 by registrars) for costs analysis. A sub-sample of 56.5 patients 7–10 days after hospital attendance and aggregate costs of hospital care provided for satisfaction and outcome analysis 4 3-month follow-up for clinical outcome analysis 5 For outcome analysis, telephone interviews and postal surveys were used. Details of the methods to derive costs (diagnostic tests, treatment, referral, doctor's time and transactions) were not stated. It was stated that the prescription costs were estimated from the hospital pharmacist's price list and the costs of doctor's time was estimated from their employment costs, converting these to costs per minute after adjusting for working hours and leave.
Results Quantitative results	Outcome and costs: Patients' reported outcome and use of general practice in 7–10 days after attendance were similar: 85%, 85% and 88% of those seen by general practitioners, senior house officers and registrars respectively were fully recovered or improving (p = 0.840), while 20%, 18% and 21% respectively consulted a general practitioner or practice nurse (p = 0.774). Excluding costs of admissions, the average costs per case were £19.30, £19.97 and £11.70 for senior house officers, registrars, and general practitioners respectively. With cost of admissions included, these costs were £58.25, £44.68 and £32.30 respectively.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No 2 No 3 No. For outcome analysis, telephone interviews were conducted. However, a postal questionnaire was sent to patients if they lacked a telephone. 4 3-month follow-up of clinical outcome by using a questionnaire completed by the patients' family practitioner. 5 The assignment of general practitioners, senior house officers or registrars was not random. The selection of patients' samples was random. 6 An inner city accident and emergency department in south east London
Commentary	Estimation of hospital costing data was not very accurate and the research team could not estimate the cost impact of the differences between the groups of doctors in their referral rates to general practice and other primary care services in the community, nor could they estimate the costs of A&E follow-up or rehabilitation. Also, they could not calculate management costs involved in administrating the scheme. Thus, there may be important hidden costs that should be considered.
Research implications	Costs to patients and their families in using the accident and emergency again rather than using their general practise will be the subject of a future study.

ID, origin, authors (year)	223, USA, Dellasega, C.A. and Zerbe, T.M. (2000)
Aims	To use quantitative data to compare the outcomes for frail 'rural elders' patients who received post-discharge care provided by advanced practice nurses (APN) with those with no nursing care, RN care only and both RN and APN care. To investigate the role of the APN in delivering post-discharge intervention for frail elderly hospital patients through focus-group interview. *Workforce:* Nurses; secondary *Feature:* Specialisation *Outcome:* Cognitive functioning, self-rated health, informal services provided and use of health care resources
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective randomised controlled study (Part I); one focus-group interview (Part II) 2 Included elders who were in one of the study-site hospitals, who were scheduled for discharge to home and frail. Excluded those who refused to continue for the study because of dissatisfaction with the research process and feeling too overwhelmed in the home care environment as well as loss of contact with the participants and death of the elder. 3 140 elders (43 with no support, 30 both RN and APN, 39 APN only, 28 RN only) and 65 caregivers (20 with no support, 17 with support from both RN and APN, 17 with APN only and 11 with RN only) for Part I; the focus group (4 APNs) interview conducted by the two authors. 4 N/A 5 Patient chart form and the Caregiver Information Form were used to collect data on demographic variables. Use of health care resources was evaluated through the Resource Utilisation Checklist. A self-rated health of elders was evaluated using one item from the 36-item Short Form Medical Outcome Survey. Caregiver well-being was measured by two components of the Caregiver Burden Inventory, the Time and Effort and Thoughts and Feelings subscales. Data were collected at baseline, discharge, and 2, 4, and 6 weeks after discharge.
Results Quantitative results	Elders in the APN-only group experienced fewer emergency room visits and hospital re-admissions, but the difference was not significant. Caregivers receiving APN-only support reported significantly fewer work days missed compared to the RN support caregivers. In the focus group, APNs perceived their role as more comprehensive than autonomous, addressed gaps in care, and focused on informing, counselling, and teaching patients and their families. They focused on both patients and their caregivers and were able to enhance continuity of care.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No 2 No 3 No. An initial participation consent rate of 50% for elders and 56% for caregivers only. Most data forms were self-reported and some participants dropped out of the research. 4 N/A 5 Yes. Patients were randomly assigned to have no support, RN only, APN only and both RN and APN supports. 6 Not stated
Commentary	A low initial consent rate because of the illness of the subject pool and the rural location of residence; relatively small sample size
Research implications	Advanced practice roles for nurses in the community are expanding rapidly. As these positions develop and evolve, research on their impact on patient outcomes needs to be conducted to validate and reinforce the unique and important contributions of APNs.

ID, origin, authors (year)	167, Canada, Di Carlo, A. *et al.* (2001)
Aims	To determine outcome differences in the management of fistulas complicating diverticulitis between patients under the care of specialists (colorectal surgeons) and general surgeons. *Workforce:* colorectal surgeons and general surgeons *Feature:* diagnostic investigations, operative findings, operative management, postoperative management *Outcome:* use of diverting procedures (e.g. colostomy), postoperative complications and length of hospital stay (LOS)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Historical cohort study/review 2 Inclusion: All cases of fistula complicating diverticular disease that were operated on in four university affiliated hospitals in Quebec between 1975 and 1995. Exclusion: 3 patients who underwent non-operative management were excluded. 3 122 patients: 37 under the care of colorectal surgeons (elective surgeries) and 85 under the care of general surgeons (84 elective, 1 semi-urgent surgery). 4 N/A 5 Review of hospital charts, electronic database of discharge summaries from 21 year period
Results Quantitative results	There were no significant differences in patient demographics, preoperative comorbidities, or the number of preoperative diagnostic investigations between the two groups. The colorectal surgeons (CS) performed more intraoperative ureteral stenting (CS 55.5% vs. general surgery (GS) 24.4%, p = 0.001). The general surgeons performed more initial diverting Hartmann's and colostomy procedures (CS 5.4% vs. GS 27%, p = 0.13). Patients in the general surgery group had longer preoperative lengths of stay (median CS 3 (range 1–28) days vs. GS 8 (range 0–29) days; p <0.001), longer postoperative lengths of stay (median CS 11 (range 5–40) days vs. GS 14 (range 2–80) days; p = 0.001) and longer total lengths of stay (median CS 14 (range 6–62) days vs. GS 24 (range 6–100) days; p <0.001). The patients in the GS group experienced a higher rate of wound infections (CS 5.4% vs. GS 12.9%) and a larger proportion of them experienced complications (CS 27% vs. GS 41.2%).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment 2 Adjustment to control for confounding factors of abscess, year of surgery, and surgeon experience 3 Uniform data collection; hospital record review 4 No patient follow-up, chart search for recurrence 5 Non-random, patients non-randomly assigned to specialist or generalist. 6 Four university teaching hospitals in Quebec
Commentary	The authors conclude that specialisation in colon and rectal surgery contributed to an improved outcome, with lower rate of diverting procedures, a shorter hospital stay and a lower rate of complications. Fistula as a complication of diverticulitis is rare, and only 20% of patients require operative intervention. This is primarily a review of patient records. Some bias may occur in results as patients in the GS group had a higher incidence of abscesses prior to surgery which might impact rate of wound infection and postoperative complications.
Research implications	Exploration of whether a clear set of clinical practice guidelines based on current literature would reduce the number of colostomies.

ID, origin, authors (year)	182, USA, DiRusso, S., Holly, C., Kamath, R., *et al.* (2001)
Aims	To asses the impact on patient outcome and hospital performance of preparing for and achieving American College of Surgeons (ACS) Level I trauma verification *Workforce:* Specialist trauma staff (medical and allied): physicians, case managers, nurse practitioners, registrars and administrative support staff; primary care *Feature:* Specialisation of workforce and skill mix *Outcome:* Mortality and morbidity (complications) and length of stay. LOS was also related to costs but those results will not be reported in this abstract.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Retrospective comparative 2 Trauma patients served at a hospital in Hudson Valley region of New York State in 1994 and 1998, and the trauma teams for each year 3 1,098 patients in 1994 and 1,658 in 1998. 4 In-hospital 5 Data came from New York State Trauma Registry project in 1994 and 1998
Results Quantitative results	Trauma system improvement as related to achieving ACS Level I verification appeared to have a positive impact on survival and patient care. Mortality in 1994 was higher than in 1998. *Unadjusted risk mortality* **1994 (n=1093)** **1998 (n=1676)** ***p*-value** Mortality (total) 8.2% 6.1% <0.05 Deaths in the ER 3.1% 1.2% <0.05 Deaths (patient ISS >30) 45% 26% <0.05 *ISS= injury severity score)* *Comparison of mortality rates for 1194 and 1998 using risk models* **No. of survivors greater than predicted** **Z-statistic** **W-statistic** *TRISS-based mortality* 1994 2.36 0.43 (NS) ND 1998 12.54 2.02 0.9 *ANN based mortality* 1998 27 3.44 1.61 ROC Az= 0.93 Lemeshow-Hosmer C-Statistic 62.5 1.6 more patients per 100 trauma admissions survived in 1998. (NS = not significant; ND = not determined; ROC= receiver operating characteristic)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Using ANN or an artificial neural network mortalities were adjusted for heart rate, systolic pressure, temperature, GCS, respiratory rate haematocrit, time to emergency room, certification level responder ICD-9-CM E- code, ISS, age, sex, race, and intubation status. 2 MTOS TRISS was also used to compare risk-adjusted mortality rates, but it could not include entire sample in its analysis because of missing data. 3 Yes 4 Yes 5 No 6 Results apply to the aggregate trauma patients in the study hospital and the trauma teams and staff working there in 1994 and 1998.

Commentary	All of the improvements in care cannot be attributed solely to the increase in commitment and resources devoted to trauma care. Many other variables may have contributed such as an improvement in technology or perceptions that more is expected of these trauma teams. In-hospital mortality may not be the best indicator or performance and outcome. Mortality after discharge was not assessed.
Research implications	Study should be controlled for volume of patients as well. Is the reduction in mortality related to the increase in volume? Did patient length of stay decrease because of better management or as a result of the decrease in morbidity? A cross-sectional study is needed to ascertain if the trend is repeated across the rest of the hospitals in the Hudson Valley trauma centres. Will all states' trauma centres benefit from ACS verification? What were the differences between previous New York State criteria and the new ACS criteria?

ID, origin, authors (year)	190, UK, Dixon, D.S. *et al.* (2001)
Aims	To evaluate the effectiveness and cost implications of hospital diabetes specialist nursing care compared to non-specialist *Workforce:* Nursing; tertiary *Feature:* Specialisation *Outcome:* Length of stay, pattern of re-admission (frequency and time to first re-admission), diabetes-related quality of life, diabetes knowledge score, satisfaction with treatment, and GP and community care contacts following discharge
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective randomised controlled trial study 2 Included all diabetes patients in a single hospital during the study period. Excluded 129 out of 300 hospital Type 1 or 2 diabetes patients who were unable to complete the secondary outcome questionnaires because they were either visually impaired, non-English speaking, confused, or had reduced consciousness. A further 24 failed to return the 1-week post-discharge questionnaires, and 14 patients died before the intervention. 3 300 + 129 + 24 + 14 = 133 patients (66 control vs. 67 specialist intervention) 4 No 5 Primary outcome measures were hospital LOS, frequency of re-admission within 12 months and time in days to first re-admission. These were collected manually on all subjects and verified using the hospital patient management system. Secondary outcome: diabetes-specific quality of life, and diabetes knowledge were assessed at randomisation and at 1 week post-discharge by post, using self-completed questionnaire, the Diabetes Knowledge Scale, which was modified to provide a version for insulin users and non-users. Patient satisfaction was assessed using a modified version of the Diabetes Clinic Satisfaction Questionnaire also administrated 1 week post-discharge.
Results Quantitative results	The effectiveness and cost implications of hospital diabetes specialist nursing care was compared with non-specialist caring. Median length of stay was lower in the intervention group (11.0 vs. 8.0 days, p <0.01). Re-admission rates were the same in the two groups (25%). When the reduced length of stay was accounted for, the nurse specialist intervention produced a mean cost per admission of £436 lower than that of the control group (p = 0.19). Patients in the intervention group were more satisfied with their care.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No 2 No 3 Collection of primary outcome measures were uniform. Secondary outcome measures were not as uniform because questionnaires were self-completed. 4 Many were unable to finish the study (see Methods 2) 5 Yes, patients were randomly assigned to control group and intervention group. However, a significant portion of patients couldn't finished the questionnaires and therefore the sample size was not large. 6 a single UK university hospital
Commentary	The number of secondary outcome measures was small and therefore these data are open to bias.
Research implications	Future research looking at costs and health outcomes (including generic quality of life measures) following discharge are needed before the impact on longer-term diabetes care is fully understood.

ID, origin, authors (year)	640, USA, Frances, C,D, *et al.* (1999)
Aims	To evaluate whether cardiologists provide more recommended therapies to elderly patients with acute myocardial infarction (AMI) and, if so, to determine whether variations in processes of care account for differences in patient outcome. *Workforce:* Physician; non-federal acute care hospitals *Feature:* Specialisation *Outcome:* Percentage of 'good' and 'ideal' patients for a given AMI therapy who actually received that therapy, percentage who received stress testing or coronary angiography, revascularisation rates and 1-year mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort study 2 Identified and included 12,150 Medicare beneficiaries 65 years and older with AMI. Excluded patients who were transferred to another institution (n=2100, 17%), those with missing data (n=1164, 10%). There were then 558 patients whose physician UPIN did not match the UPIN data file, and 665 patients with an associated UPIN designating a different physician specialty and so these patients were excluded. 3 12150 – 2100 – 1164 – 558 – 665 = 7663 patients 4 1-year mortality follow-up 5 The Cooperative Cardiovascular Project collected data abstracted from the medical charts of Medicare patients in California who were discharged with AMI from an acute care hospital between April 94 and July 95.
Results Quantitative results	Treatments and outcomes in AMI patients treated by cardiologists and generalists were compared. During hospitalisation, good candidates for aspirin were more likely to receive aspirin if they were treated by cardiologists (87%) than by medical subspecialists (73%; p = 0.001), general internists (84%; p = 0.003), or family practitioners (81%; p <0.001). Cardiologists were also more likely to treat good candidates with thrombolytic therapy (51%) than were medical subspecialists (29%; p <0.001), general internists (40%; p <0.001), or family practitioners (27%; p <0.001). Similar 30-day mortality rates across physician specialties were found. However, 1-year mortality rates were greater for patients treated by medical subspecialists (OR 1.2), general internists (OR 1.1), and family practitioners (OR 1.3).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 30-day and 1-year mortality were both adjusted for patient demographic, comorbidity, and severity of illness characteristics. 2 No 3 Yes 4 Completed 1 year mortality follow-up. However, lots of missing data and thus, lots of patients were excluded from the study (see Methods 2). 5 No. Ethical and feasibility considerations limit the randomised assignment of specialists. Also, a large portion of identified participants were excluded from the study for various reasons (see Methods 2). 6 Involved all acute care hospitals in California.
Commentary	Unable to determine whether patients who were cared for by generalists received follow-up or consultation with cardiologists. Excluding transferred patients, who may differ from non-transferred patients, may bias mortality and utilisation rates by physician specialty.
Research implications	With the exception of the in-hospital use of aspirin, recommended AMI therapies were markedly underused, regardless of specialty of the physician. Policies should be aimed at improving the care provided by all physicians.

ID, origin, authors (year)	701, USA, Franks, P. and Fiscella, K. (1998)
Aims	To examine through a patient-focused-orientation whether patients using a primary care physician have lower expenditures and mortality than those using a specialist. *Workforce:* Personal care physician (general practitioner, family physician, internist or obstetrician–gynaecologist) or specialist *Feature:* Medical diagnoses and patient care *Outcome:* Total annual health care expenditure and 5-year mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective cohort study 2 Inclusion: Respondents 25 and older who reported using one or more physicians as usual source of care. Exclusion: Patients without complete medical records. 3 13,270 (95.9% of eligible) patients. Of these 12,213 had a personal care physician. 4 5-year mortality verified through National Death Index. 5 During the 1987 National Medical Expenditure Survey (NEMS) 4 interviews were performed to collect data on medical care, health experiences, health insurance, and a subjective check list of health status.
Results Quantitative results	Respondents with a primary care physician as a personal physician were more likely to be women, white, rural, report fewer medical diagnoses and higher health perceptions and have lower annual health care expenditures (mean: $2029 vs. $3100) and lower mortality (hazard ratio = 0.76, 95% CI 0.64–0.90). After adjustment for demographics, health insurance status, reported diagnoses, health perceptions and smoking status, respondents reporting using a primary care physician compared with those using a specialist had 33% lower annual adjusted health care expenditures and lower adjusted mortality (hazard ratio = 0.81; 95% CI 0.66–0.98). The findings provide evidence for the cost-effective role of primary care physicians in the health care system.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjustment for demographics, health insurance status, reported diagnoses, health perceptions and smoking status 2 Weights used on public-use tapes to adjust for over sampling and non-response bias. Adjustment for expenditures and adjustment for hazard ratio. 3 Uniform data, nationally representative sample of adult respondents to the 1987 National Medical Expenditure Survey 4 N/A 5 Not random 6 Nationally representative
Commentary	Findings support the idea that primary care adds value to the health care system. This is a population-based study; the authors are trying to extend the study, because patients do not present to primary care physicians with disease-specific concerns, but with 'sundry complaints'. By confining analysis to patients for whom a diagnosis has already been made, outcomes are missed for patients who present with a similar problem but who have had a different diagnosis or none. Limitations: There was no control for illness severity, and difference in case mix between physician groups was based on self-report. Self-report morbidity may have introduced a bias that significantly underestimates differences in patient morbidity between the two physician groups. An obstetrician–gynaecologist would probably be considered a specialist by most of the general population. Subjects who reported that they 'saw different doctors' were included with the specialist group. The term 'specialist' could have been more clearly defined.
Research implications	More research is needed on how to optimally integrate primary and specialty care. Evaluation of the important dimension of co-ordination of care between specialists and personal care physicians. Research on the potential benefit of the gatekeeping function of primary care physicians.

ID, origin, authors (year)	177, USA, Gillum, L.A. and Johnston, S.C. (2001)
Aims	To evaluate, through outcomes-based research, the efficiency of types of stroke centres in attaining minimal published criteria. *Workforce:* Institutions with acute stroke centres *Feature:* Attending neurologist, written care protocols, emergency medical services availability *Outcome:* In-hospital mortality and functionality outcomes, LOS, total hospital charges
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective cohort study supplemented by questionnaire-based survey 2 32 of 42 academic medical centres in the University Health Systems Consortium responded to a questionnaire detailing stroke treatment practices. Of these 29 were included in discharge abstracts in the UHSC database. 3 10,880 admissions for ischemic stroke at 29 institutions 4 N/A 5 Database of discharge abstracts, medical record review of cases between June and December 1999. Questionnaires relating to stroke treatment were completed by a hospital administrator or a specialist at each centre.
Results Quantitative results	32 institutions completed the questionnaire, and 29 of these were included in the database of discharge abstracts. In-hospital deaths occurred in 758 (7.0%) of the 10,880 ischemic stroke patients admitted through the emergency department. In-hospital deaths were less frequent at hospitals with a vascular neurologist (OR 0.51; 95%CI 0.36– 0.74; p <0.0001) and at those with guidelines stating that only neurologists could administer tPA (OR 0.65; 95%CI 0.49–0.88; p = 0.004). There was a trend toward fewer deaths at hospitals with a dedicated stroke team available by pager (OR 0.76; 95% CI 0.56–1.04; p =0.09). The presence of a dedicated neurological intensive care unit, stroke unit, and written clinical pathway for stroke were not significantly associated with in-hospital death. LOS was shorter at hospitals with a vascular neurologist (p = 0.01) Academic medical centres with a vascular neurologist and those with written guidelines limiting tPA administration to neurologists had lower rates of in-hospital mortality for ischemic stroke patients. There was a trend toward fewer deaths at hospitals with a dedicated stroke team available by pager. LOS was shorter at hospitals with a vascular neurologist.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Case mix adjustment not specified 2 None specified 3 Two types of data: administrative data supplemented by detailed medical record review and unit specific questionnaire 4 N/A 5 Not random 6 Geographic dispersal. 29 institutions, of one University Health Systems Consortium
Commentary	Limitations: lack of clear definition of ‘stroke team’.
Research implications	Further study which includes non-academic medical centres would validate these findings.

ID, Origin, Authors & Year	1117, Italy, Grilli, R., Minozzi, S., Tinazzi, A. *et al.* (1998)
Aims	To determine whether cancer patients receive more appropriate diagnostic and therapeutic interventions or have better outcomes when cared for by specialised centres/clinicians. *Workforce:* Mixed workforce and settings *Feature:* Specialisation (individual clinicians, specialisation of institutions (hospital, centres) and proxy indicators including hospital teaching status and hospital size) *Outcome:* Mortality at 3 and 5 years, proportion of patients treated according to optimal care criteria, loss to follow-up, or having defined investigative procedures, proportion with complete information on staging, histology, use of breast conserving surgery or specified cancer care management including pain management and number of surgical interventions required
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 Individual study design (n): Randomised controlled clinical trials (RCTs) and prospective and retrospective cohort studies that compared clinicians or centres grouped according to the definition of specialisation were included if they were published since 1980 and provided information about the association between objective measures of process or outcome of care for cancer patients and degree of specialisation. Oncology patients with the following types of cancer were studied: cervical; breast; ovarian; Wilm's tumour; rhabdomiosarcoma; medulloblastoma; haematological; and various other cancers including colorectal, lung, prostate, and testicular. 3 – 4 47 papers describing 46 studies included in the final analysis. 5 MEDLINE and EMBASE were searched for articles published in the English language between 1980 and 1995. Details of terms used are given. Reference lists of papers identified and review articles were also examined 6 The following adjustment score was applied based on the number and type of patient characteristics considered in adjusting for case mix: score 0 awarded for no case adjustment; score 1 awarded for adjustment for demographic variables or only for stage of disease; score 2 awarded for adjustment for demographic variables, comorbidity, or for demographic variables and stage of disease; and score 3 awarded for demographic variables, comorbidity and stage of disease. 7 Investigation of differences and bias: Results from different studies were only pooled when the impact of specialisation was assessed on the same tumour, the comparison was sufficiently adjusted for case mix (adjustment score >2), and a common outcome (3- or 5-year mortality) was used. Summary OR and 95% CI were estimated from the mean weighted by the reciprocal of the variance. Differences were discussed. The OR with 95% confidence intervals for the effect of specialisation on various outcomes were presented graphically. The chi-squared statistic was used to estimate statistical heterogeneity for mortality in studies on breast cancer and studies on ovarian cancer.

Results Quantitative results	*Specialisation and process of care:* 11 observational studies provided information on the impact of specialisation for various cancer sites. 5 defined specialisation at the clinician level and 6 at the level of centres. Overall results favoured specialised clinical centres. Only 5 studies adjusted adequately for the case mix between comparison groups. Studies were mostly low quality and tended to show cancer centres performed specific diagnostic staging procedures more often in breast cancer, childhood cancers and ovarian cancers. Breast conserving surgery (3 studies) was more frequently offered in centres with oncology departments or wards. *Proxy definitions of specialisation and process of care:* 17 studies compared hospital patterns of care according to teaching status (11 studies) and hospital size (5 studies). 13 studies were on breast cancer, 2 on ovarian cancer or included multiple sites. Studies scoring 2 or more on case mix adjustment criteria showed greater reporting of clinical and pathological staging in the notes and greater use of two-stage surgery in larger or teaching centres. Conservative surgical procedures were more commonly used in larger or teaching centres. No difference between non-specialised vs. specialised was noted in the use of adjuvant chemotherapy for breast cancer. *Specialisation (however defined) and mortality:* Generally patients had a lower risk of long-term mortality when treated by specialised centres/clinicians though results from two studies differed. *Specialisation (however defined) and mortality for breast cancer (5 studies):* All had an adjustment score of 2 or more. Lower 5-year mortality reported when treated in specialist centres or by specialised clinicians OR = 0.82 (95% CI 0.77, 0.88). Heterogeneity chi-squared = 0.08 (p = 0.99). *Specialisation (however defined) and mortality for haematological cancer (4 studies one of which dealt with 3 types of tumour, giving 6 treatment arms):* 5 of the 6 treatment arms showed lower mortality when treated in specialised situations. *Specialisation (however defined) and mortality for ovarian cancer (7 studies):* 6 of 7 studies showed lower mortality when treated in specialised situations. Quality of studies and definition of specialisation differed. Heterogeneity chi-squared = 4.5, p = 0.60. *Specialisation and mortality for other solid tumours (5 studies):* 2 studies reported statistically significantly lower mortality for colorectal cancer and prostate cancer in teaching vs. non-teaching hospitals. Lung cancer (1 study, 2 histological types): results differed according to histology. Testicular cancer (1 study): showed an advantage only for the availability of on-staff urologists and not for oncologist. Few studies focused on types of neurological tumours, sarcomas, or childhood cancers. There were only a limited number of poor-quality studies in these fields. *Impact of specialisation on outcomes other than long-term mortality. Quality of life in breast cancer (1 RCT):* No difference between groups. Studies reporting postoperative/in-hospital mortality in gastrointestinal (1 study), lung (1 study) and ovarian (1 study) showed contradictory results. Despite the fact that care provided by specialised centres/clinicians appeared to be better, both when assessed in relation to process indicators and to mortality, this evidence should be considered far from conclusive because of major methodological flaws in these studies.
Commentary	The authors do not state how the papers were selected for the review, or how many of the reviewers performed the selection. Although validity was assessed the authors do not state how papers were assessed for validity, or how many of the reviewers performed the assessment. No details are given of methods used to extract data. The aims and inclusion criteria were clearly defined. There was discussion of the possibility of publication bias due to outcome-dependent publication and the influence of methodological flaws on the results: lack of comparability of patients seen at specialised and non-specialised centres; use of observational studies resulting in an over-estimate of effect size; and the unknown clinical relevance of diagnostic or therapeutic procedures to patient outcome. By limiting primary studies to those published in English some relevant studies may have been omitted. The authors correctly advise caution in interpretation of the results in the light of the overall limited quality of evidence identified.
Research implications	The authors consider that little effort has been made to disentangle key components of specialised care and that the question of whether non-specialised providers could achieve the same clinical results as their specialised colleagues when adequately trained has been poorly addressed by available research.

ID, origin, authors (year)	396, UK, Hearn J and Higginson IJ (1998)
Aims	To determine whether teams providing specialist palliative care improve the health outcomes of patients with advanced cancer and their families or carers when compared to conventional services *Workforce:* Specialist palliative care teams (Doctors, clinical nurse specialists, social workers, chaplains, therapists and psychologists or psychiatrists), Mixed *Feature:* Specialisation *Outcome:* Patient satisfaction, the patient being cared for where they wished, family satisfaction, family anxiety, patient pain and symptom control
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 *Inclusion:* Any type of study which considered the use of specialist teams caring for advanced cancer patients and their families was included. *Exclusion:* Studies focusing on one cancer site. 3 Randomised controlled trials (5) involving 925 patients and 344 carers 4 Observational or comparative studies (13) involving 14,466 patients and 577 carers 5 Databases searched: Medline (1982–1996), PsycINFO (1984–1996), Cinahl (1982–1996), Bids (1992–1996), Embase (1992–1996), Social SciSearch (1992–1996) and IBSS (1992–1996). *Palliative Medicine*, the *Journal of Palliative Care* and *Progress in Palliative Care* were hand searched from their first issue to 1996. Internet sites: CancerWEB and OncoLink. Authors of ongoing trials were contacted – identified from conference proceedings, by searching references from seminal articles and through collaboration with researchers conducting related reviews. 6 A grading system was used to evaluate the validity of the primary studies (I-IV). The appropriateness of the various outcome measures used were taken into account when allocating a grade to each study. 7 Investigation of differences and bias. The studies were not combined. Some study differences were discussed in the results and discussion sections of the review.
Results Quantitative results	When specialist multiprofessional care was compared with conventional care, 4 of the 5 RCTs and the majority of the comparative studies indicated that the specialist, co-ordinated approach resulted in similar or improved outcomes.
Commentary	The authors do not state how the papers were assessed for validity, or how many of the authors performed the validity assessment. Although study validity was assessed by a grading system, the results and conclusions did not take these into account. Results in the tables were summarised in words and no actual data reported other than costs. The majority of the tabulated primary data were provided; however, a few did not provide information such as the age of the patients. The authors' conclusions may overstep the quality of the data presented.
Research implications	High-quality research into the costs and benefits of multiprofessional teams for palliative care should be designed and carried out.

ID, origin, authors (year)	354, USA, Jackson, J.L. *et al.* (1999)
Aims	To explore the effect of interns' involvement on patient care outcomes in a walk-in general medical clinic *Workforce:* Interns and staff physicians *Feature:* Symptom-related expectations of patients and functional status. Physician perceptions of difficulty of patient. *Outcome:* Symptom outcomes and satisfaction, illness worry and unmet expectations
Methods	
1 Design	1 Pilot project questionnaire-based study
2 In-/exclusion	2 Convenience sample of ambulatory patients in a walk-in setting; no information of inclusion or exclusion
3 Sample size	3 750 patients: 195 seen by interns, 555 seen by staff physician; 28 interns and 26 staff physicians
4 Follow-up time	4 Immediately after visit, 2-week and 3-month follow-up
5 Data collection: source and period	5 Questionnaire surveys filled pre-visit, and at 3 follow-up points. Physicians were surveyed about perceptions about patients; between January 1995 and August 1998 at the Walter Reed Army Medical Center in Washington DC
Results Quantitative results	Quantitative results showed there were no differences in visit costs, subspecialist referrals, health utilisation or hospitalisation rates. Patients of both interns and staff physicians experienced the same reduction in serious illness worry immediately after their visits (64% vs. 18%, p <0.001), few patents in either group had unmet expectations after the visit. Most of the patients' conditions had improved by two weeks (53%) and three months (79%), with concomitant decreases in symptom severity, not different between the groups.
Quality appraisal	
1 Case mix adjustment	1 No case mix adjustment
2 Other adjustment	2 No other adjustment
3 Uniform data collection	3 2 types of interview questions and data to patients, one questionnaire completed by attending physicians and interns
4 Participant follow-up	4 Follow-up by survey at 2 weeks (690 completions) and 3 months (612 completions)
5 Random sampling	5 No randomisation, convenience sample assigned to attending on first come, first serve basis.
6 Geographical dispersal	6 One walk-in clinic at a military hospital
Commentary	Qualitative results showed comparative post-visit satisfaction, residual expectation, symptom resolution and functional status improvement. Interns were allotted a longer period for appointments and spent more time with patients. This may have impacted qualitative reports. Limitations: Single clinic at single military site.
Research implications	Further study of residents providing independent care in ambulatory settings.

ID, origin, authors (year)	91, UK, Jarman, B. *et al.* (2002)
Aims	To determine the effects of community-based nurses specialising in Parkinson's disease on health outcomes and health care costs *Workforce:* Nurses; primary *Feature:* Specialisation *Outcome:* Survival, stand-up test, dot in square test, bone fracture, global health question, PDQ-39, Euroqol and healthcare costs
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Randomised controlled trial 2 Sampling frame included all English health authorities that did not already have well-developed community-based services of nurse specialists in Parkinson's disease. Eligible patients were those taking one or more anti-Parkinsonian drugs. Excluded patients aged 17 years or less or those with severe mental illness or cognitive impairment sufficient to preclude valid informed consent. 3 1859 patients with Parkinson's disease agreed to participate; of those 23 died before start of intervention, 1028 participated in the nurse specialist group and 808 patients participated in the control group. 4 Follow-up of mortality continued for 4 years. 5 Mortality data from NHS Central Registry; other clinical outcomes (stand-up test, proportion sustaining fracture) and health care costs by face-to-face patient interview; patient well-being by self-completed questionnaire.
Results Quantitative results	The effects of community-based nurses specialising in Parkinson's disease on health outcomes and costs were examined. After 2 years, 315 patients (17.3%) had died. Mortality did not differ between those who were attended by nurse specialists and those receiving standard care from their general practitioner (hazard ratio for nurse group vs. control group 0.91, 95% confidence interval 0.73 to 1.13). Also, nurse specialists had little effect on other clinical outcomes of Parkinson's patients. However, scores on the global health question (patients' sense of well-being) were significantly better in patients attended by nurse specialists than in control. For health care costs, there are no differences between the two groups.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Mortality measure was not adjusted with acuity. However, for stand-up test score, three groups were categorised (no problems, without holding on and unable/had to hold on) to give a fair comparison between the groups at the end of the study. 2 No 3 Uniform data collection 4 Completed follow-up of mortality . 5 Random sampling was performed by an independent social survey organisation. Nine health authority areas were randomly selected and all the general practices in the nine areas were approached. 1859 patients were identified as Parkinsons' patients. 56% were randomly assigned to the nurse specialist group and 44% to the control group. 6 Patients spread all over in UK.
Commentary	The trial intervention used nurses who had only recently trained in nursing patients with Parkinson's disease. They were therefore on a professional learning curve and may not be representative of experienced nurse specialists.
Research implications	There was a significant improvement in subjective well-being of patients cared for by a nurse specialist. This improvement was achieved without an increase in health care costs.

ID, origin, authors (year)	714, USA, Jollis, J.G. *et al.* (1996)
Aims	To examine mortality of patients hospitalised for myocardial infarction (MI), according to the specialty of the admitting physician *Workforce:* Primary care physicians and specialists *Feature:* Resource-intensive care (use of medications and coronary revascularisation) *Outcome:* Use of specified drug therapies, cardiac procedures, LOS and survival
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort design 2 Inclusion: Of 220,535 Medicare patients over the age of 65 who were hospitalised for acute MI in 4 states, a subgroup of 8241 were identified for whom there were Cooperative Cardiovascular Project (CCP) data. Exclusion: Patients who were receiving therapy before admission were excluded in an adjustment for medication use, and patients who underwent coronary revascularisation before discharge excluded to adjust for length of hospital stay. 3 CPP cohort of 8241 4 In-hospital, 30-day and 1-year survival rate 5 Clinical data from two databases: Cooperative Cardiovascular Project (CCP) and the Medicare National Claims History File from a 7-month period from June to December of 1992
Results Quantitative results	Patients admitted by cardiologists were 12% less likely to die within 1 year than those admitted by a primary care physician. After adjustment for patients' characteristics, patients admitted by cardiologists had significantly better 1-year survival than those admitted by physicians in all the primary care specialties (hazard ratio, 0.87; p <0.00.). The survival advantage for cardiology persisted and remained significant after adjustment for hospitals' characteristics (hazard ratio 0.88; p <0.001). Adjusted rates of 1-year survival did not differ significantly among the primary care specialties. Cardiologists had the highest rate of use of cardiac procedures and medications including medications that are associated with improved survival.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Case mix adjustment for patient characteristics 2 Adjustment for hospital characteristics 3 Uniform data collection from two sources, Medicare admission records, abstracted data from CCP database 4 Retrospective follow-up over a 7-month period 5 No randomisation 6 Four states: Alabama, Connecticut, Iowa and Wisconsin
Commentary	Specific aspects of care by cardiologist (narrower clinical focus) and admission to a hospital that cares for large numbers of patients with MI, initiation of early treatment, and on-site availability of procedures for management of complications are two factors whose influence is difficult to separate. Results should be generalisable to more recent cohorts of this age group. Results indicate a critical need to define better the difference between specialty and primary care and the effects of those differences on outcomes.
Research implications	Further research that further assesses use of current protocols and drug therapy use by personal care physicians to determine whether standardised practice improves outcomes.

ID, origin, authors (year)	255, USA, Kenyon, T.A.G., Lenker, M.P., Bax, T.W. and Swanstrom L.L. (1997)
Aims	To determine if the presence of a well-organsed, dedicated laparoscopic OR team will improve surgical outcomes for this procedure *Workforce:* Specialist and doctors and nurses, OR technicians and support staff; secondary care *Feature:* Specialisation and training of workforce *Outcome:* Length of surgery (length of anaesthesia time) and complications (conversion to open procedures and major complications)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective observational 2 The hospital records of patients who underwent uncomplicated laparoscopic cholecystectomy (LC) from 1990 through 1993 and the experience of two LC surgeons: advanced LC surgeon (>200 LC cases and routinely performing other advanced procedures) and basic LC surgeon (<50 LC cases and not performing advanced procedure). Data were collected only on LC with intraoperative cholangiograms from two sites. Operations that involved more than a LC with cholangiograms or those listed as emergent were not evaluated; procedures that converted to open were not included in the overall OR time results. 3 71 patient cases: 27 with dedicated team, 44 with the non-dedicated team 4 Complete 5 Data came from hospital records dated 1990 through 1993.
Results Quantitative results	The designated laparoscopic team had decreased operative time, and a smaller conversion rate for the less-experienced surgeon. Mean anesthesia time for basic surgeon: alpha/beta: 144.2 (± 11.7) / 175.7 (± 5.7) minutes; p <0.05 Mean anesthesia time for advanced surgeon: 97.5 (± 6.3)/ 128.9 (± 7.7) minutes; p <0.05 Alpha site, designated trained laparoscopic team; beta site, randomly assigned OR team: Y surgeon, basic laparoscopic surgeon; X surgeon, advanced laparoscopic surgeon. *Laparoscopic to open conversion rates between sites alpha and beta, and surgeon X and surgeon Y* Site: cases (%) alpha/beta: 0/27 (0); 4/40 (10) Surgeon: cases (%)X/Y: 0/35 (0); 4/36 (9) There were no cases by either surgeon converted to open at the alpha site. There was, however a 9% rate of conversion for all cases done at the beta site. All conversion cases done by the basic surgeon. The surgeon's conversion rate was therefore 14% at the beta site. In both groups there were no major operative complications.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Patients were controlled for age, gender, and ASA class but differences were not significant. 2 Case characteristics were matched between surgeons. 3 Yes 4 Yes 5 No 6 Results apply to laparoscopic cholecystecotmy patients at the two institutions
Commentary	Does not state location of study. Small sample size. Study failed to report any limitations. A trained laparoscopic team is able to handle the advanced equipment more efficiently which decreases patient downtime, accidental injuries to patients and staff, and allows surgeons to focus on the operation which results in fewer delays and shorter operative times.
Research implications	Study should be done with a larger sample size to confirm results. Does the amount of cases handled by teams act as a confounding factor to improvement of outcomes with more specialised team?

ID, origin, authors (year)	1162, USA, Krapohl, G.L. *et al.* 1996
Aims	To explore the increased use of unlicensed assistive personnel in nursing care delivery in the acute care setting. *Workforce:* Unlicensed assistive nursing personnel (UAP), RNs *Feature:* Cost and productivity related to level of staffing, patient case mix (length of stay and acuity), customer expectations, variation in nursing supply and demand (move to independent and collaborative practice in primary care and community settings) *Outcome:* Quality of care, patient satisfaction, nurse satisfaction, cost/productivity and efficiency
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Literature review 2 Not reported 3 Not reported 4 Research studies, evaluative reports and case studies which were found through search with subject headings: assistive personnel, nurse extender, support personnel, and unlicensed assistive personnel. Total (n)=19 5 CINAHL (1988–1994) 6 No validity criteria 7 Investigation of differences and bias: data tabled as Clinical models, Integrated models, and Non-clinical models
Results Quantitative results	No empirically strong evidence was found to confirm that nursing support personnel improved quality or increased nurse and patient satisfaction. Two major areas of concern regarding the use of UAP were identified in the literature: personnel-related problems and RN preparation.
Commentary	This review provides basic, tabled detail regarding whether the studies dealt with nurse and patient satisfaction, cost and quality. A short paragraph summarising results and commentary as to the length of the study and how factors were studied. The authors note that studies were frequently anecdotal in nature, were conducted at a single institution, lacked comparison groups, used instruments of untested reliability and validity, and were of small sample size. No empirically strong evidence was found to confirm that nursing support personnel improved quality or increased nurse and patient satisfaction. Nor were these studies sufficiently rigorous in nature to measure costs since a number of multiple variables such as increased supervisory personnel, re-admission rates, length of stay, and training costs were not addressed. All studies summarised, no weighting. Simply a literature review. No information as to the number of reviewers and method of assessment.
Research implications	Recommendations from the review were to: • improve empirical evaluation, systematic and comparative evaluation of quality of care, patient and nurse satisfaction and costs/productivity • examine the impact of UAP as they affect the entire system (e.g. costs may simply be redistributed rather than reduced). Improvement of technical support systems can improve efficiency and effectiveness of nursing care delivery.

ID, origin, authors (year)	39, UK and Netherlands, Laheij, R., van Marrewijk, C., Buth, J. *et al*. (2002)
Aims	To determine whether the experience of the specialist team was associated with adverse events following endovascular treatment of abdominal aortic aneurysms (AAA) *Workforce:* Specialist physicians; secondary care *Feature:* Training of workforce, i.e. experience, level of skill *Outcome:* Patient mortality and adverse events measured by the relative risk of death and the need of secondary intervention among patients who underwent endovascular stenting, according to the experience of the specialist teams
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional, observational 2 It looked at patients who underwent endovascular abdominal aortic aneurysm repair between January 1994 and July 2000. Exclusion was done after analysis and removed 91 patients by consent. 20 of those patients had been admitted recently and would have been operated on in the near future. 3 2863 patients from 93 hospitals, in 16 European countries 4 Follow-up took place at 1, 3, 6, 12 and 18 months and then yearly for a total of 4 years inclusive. 5 Data were collected from the European Collaboration on Stent/graft aortic Aneurysm Repair (EUROSTAR) database on patients who underwent endovascular AAA with in the time period.
Results Quantitative results	Mortality and adverse events leading to secondary intervention after endovascular AAA repair were significantly lower in patients who underwent endovascular AAA repair by a highly experienced specialist team of vascular surgeons and intervention radiologists than in those who underwent endovascular stenting by a relatively inexperienced team. *Relative risk of death and need of secondary intervention among patients who underwent endovascular stneting, according to the quartile of experience of the specialist team*

		Hazard ratio (95% CI)	
Model	***Quartile 2***	***Quartile 3***	***Quartile 4***
Death			
Unadjusted	1.37 (0.8–1.5)	0.76 (0.5–1.1)	0.64 (0.4–1.0)
Adjusted for demographic characteristics	1.13 (0.8–1.5)	0.80 (0.6–1.2)	0.64 (0.4–1.0)
Adjusted for demographic and clinical characteristics	1.11 (0.8–1.5)	0.78 (0.5–1.1)	0.70 (0.5–1.1)
Adjusted for demographic and clinical characteristics and vascular morphology	1.07 (0.8–1.5)	0.71 (0.5–1.0)	0.69 (0.4–1.1)
Adjusted for demographic and clinical characteristics, vascular morphology and endograft characteristics	1.03 (0.7–1.4)	0.61 (0.4–.0.9)	0.60 (0.4–1.0)
Secondary interventions			
Unadjusted	0.85 (0.7–1.1)	0.58 (0.4–0.8)	0.39 (0.3–0.6)
Adjusted for demographic characteristics	0.78 (0.6–1.0)	0.56 (0.4–0.7)	0.36 (0.3–0.5)
Adjusted for demographic and clinical characteristics	0.79 (0.6–1.0)	0.55 (0.4–0.7)	0.35 (0.3–0.5)
Adjusted for demographic and clinical characteristics and vascular morphology	0.55 (0.4–0.7)	0.32 (0.2–0.5)	0.78 (0.6–1.0)
Adjusted for demographic and clinical characteristics, vascular morphology and endograft characteristics	0.80 (0.6–1.0)	0.53 (0.4–0.7)	0.32 (0.2–0.5)

First quartile is reference group (hazard ratio = 1). Postoperative mortality and secondary interventions were, respectively, 40% and 68% lower in patients treated by the most experienced specialist teams compared with the least. This equals a difference in mortality and secondary intervention rates between these quartiles of 3.6 death and 9.7 secondary interventions per 100 patients respectively.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Quartiles were adjusted demographically (age and smoking status), clinically (unfit for open surgery and previous laparotomy), vascular morphology (aortic neck angulations, aortic neck and aneurysm diameter), and endograftically (device type and configuration). Team experience was based on number of patients operated on. First quartile 1–11, second 12–37, third 37–91, and fourth 92 and higher. 2 Adjustments were made for differences in length of follow-up periods between the quartiles. This was corrected by estimating the probabilities of survival and freedom of secondary interventions using the Kaplan-Meier method. In case more than 5% of the data were missing, a dummy variable was added to the model. 3 Yes 4 1412 patients were followed for at least 1 year, 632 for 2 years, 235 for 3, and 85 patients for 4 years. 5 No 6 This study represents endovascular AAA patients from 1994–2002 in 16 countries in Europe.
Commentary	Aneurysm rupture was also measured as an outcome; however, there were too few ruptures to permit conclusions to be drawn. The study did not control for any variations within the surgical teams or differences in hospitals. Limited data source.
Research implications	How many patients must be operated on to achieve the lower mortality rates and secondary interventions? How much more training is needed? Does the type of hospital influence (e.g. teaching or high technology) the occurrence of adverse events after endovascular AAA repair? Study should be repeated using data from other sources to make it more generalisable.

ID, origin, authors (year)	289, USA, McGann, P. *et al.* (1995)
Aims	To analyse the top 10 admission diagnostic-related groups in patients 65 and older to assess whether differences in quality and cost of medical care provided is influenced by specialty training *Workforce:* General and special internists vs. general and family physicians *Feature:* Resource utilisation rates *Outcome:* Morbidity, mortality, LOS, hospital charges. (Cost is defined as LOS and total charges.)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Retrospective cohort study 2 Inclusion: Hospitals in Pennsylvania that review admissions from all departments. Exclusion: Patient transfers from other acute care hospital or other site. 3 31,321 hospital admissions: 19,154 cases managed by internists, 12,167 cases managed by family physicians. 4 No post-discharge data, chart review at 8-day point 5 Data collection from the 1989 MedisGroups Comparative Database of Pennsylvania hospital admissions
Results Quantitative results	Admission diagnoses were similar for patients of family physicians and internists. After adjusting for relevant patient and hospital characteristics, there were no differences in mortality or hospital charges; however, the patients of internists experienced slightly higher morbidity (odds ratio = 1.07, 95% CI 1.017 to 1.123) and longer mean length of stay (10.80 vs. 10.54 days, p <0.05). This study suggests that it makes little difference in medical outcome or hospital charges whether family physicians or internists manage the hospital care of elderly patients for common medical problems. Savings to the health care system attributable to physician specialty may occur predominantly outside the hospital.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Case mix adjustment: controlled for patient characteristics of age, socioeconomic status, sex, admission form care facility, DRG, admission severity score. 2 Adjustment for hospital characteristics of size, type, occupancy rate, payroll expenses, availability of procedures and technology. 3 One data set, abstracts of all participating hospitals, retrospective chart review to calculate admission severity score at 3 days after admission 4 N/A 5 No randomisation 6 29 participating hospitals in one state
Commentary	Clearly defined cost analysis. Limitations: With such a large number of cases, differences that are statistically significant may not be clinically significant. Morbidity and major morbidity collapsed into single category. No race information. Potential confounding through self-selection by patient. Information concerning physician charges, re-admission rates and post-discharge mortality not available. Inability to differentiate between differences in residency training and small geographical area limits ability to generalise results.
Research implications	Specialisation does not impact outcomes significantly.

ID, origin, authors (year)	418, USA, Miller, S.K. (1997)
Aims	To explore alternative care models; the use of nurse practitioners in the hospital care of nursing home elderly result in cost-effective, quality care *Workforce:* Gerontological nurse practitioner and hospital-employed physician's assistants, both managed by physicians *Feature:* Specialty geriatric training, holistic nursing perspective *Outcome:* More effective interactions with patients and families, recognition of early indicators of change or deterioration more common in this population. Discharge and long-term needs anticipated from gerontological perspective. Continuity of care maintained.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Retrospective analysis of available data 2 Inclusion: Nursing home elderly admitted by nonteaching attending physicians (no specific numbers given) 3 284 geriatric patients admitted in 1993, dealt with by physician assistants, and 543 geriatric patients admitted in 1994, comanaged by nurse practitioner and physicians. 4 N/A 5 Anecdotal detail for 2-year period
Results Quantitative results	Mean decrease in length of stay in 20 most common diagnostic categories over the 2 years (which reflected a 1-year period without the nurse practitioner and a 1-year period with the nurse practitioner) was 2.78 days. An independent sample -test analysis demonstrates this decrease to be significant at $p < 0.05$.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No case mix adjustment 2 No other adjustment 3 Uniform collection of data, general database and anecdotal data regarding patient/family satisfaction 4 No follow-up 5 No random sampling 6 Nonteaching general medicine service in a Philadelphia teaching hospital for elderly patients from nursing care homes
Commentary	Descriptive paper of one nurse practitioner in one hospital centre. Service set up to accommodate patient population when it was discovered that the usual teaching hospital methods were not appropriate for elderly patients. Costs not documented, extrapolated from reduced LOS.
Research implications	Research is needed to document advanced practices that decrease length of stay and patient quality of life. Validate use of physician–nurse practitioner practice protocols. Research that controlled for procedure changes, transfer to intensive care, and lack of nursing home bed availability would impact on LOS outcomes.

ID, origin, authors (year)	128, Australia, Morrison, A.L., Beckman, U., Durie, M. *et al.* (2001)
Aims	To identify incidents associated with nursing staff inexperience (NSI) and estimate their effect on the quality of patient care. *Workforce:* Nurses; secondary care *Feature:* Training of workforce *Outcome:* Complications and adverse events (incidents: any unintended event that could have or did reduce the safety margin for the patient); patient satisfaction
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective 2 1472 patient incidents and nurses from ICUs submitted to the national database from the AIMS-ICU whose records showed NSI (inexperience in intensive nursing care and/or with specific procedures and equipment) 3 735 reports; 688 involved individual patients. 4 Complete 5 Data came from Australian Incident Monitoring Study in Intensive Care Units (AIMS-ICU) national database from 1993 to December 1999.
Results Quantitative results	NSI can have a negative impact on the quality of care delivered to critically ill patients as shown by the occurrence and outcome of incidents related to such inexperience. Errors are more likely to occur when NSI is combined with staff shortage, inadequate supervision and high unit activity. In 80% of all incidents the event did not cause any significant adverse outcome for the patient. Staff member precipitating incident: n=735, ICU trained/ICU not trained 35%/65% Staff member detected incident: n=735, ICU trained/ICU not trained 80%/20% Factors limiting incident effects: n=735, skilled assistance (34%), prior experience (36%), supervision (28%), rechecking patient (40%), rechecking equipment (38%) and use of protocol (26%) *Patient outcome due to incident* **Outcome** — **Individual patient reports (n=688)** Nil adverse effect — 300 Minor physiological complications — 201 Major physiological complications — 95 Physical injury — 15 Psychological injury — 11 Patient relative dissatisfaction — 24 Prolong hospital stay — 6 Unknown — 75 The reporter could select multiple patient outcomes for each report.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No adjustments were made 2 – 3 Yes 4 Yes 5 No 6 Results apply to all ICU patients from the AIM-ICU database that had incidents due to NSI from 93 ICUs.

Commentary	Data mainly just reported the incidents and their causes but did not thoroughly compare NSI incidents to those incidents with a more experienced nursing staff. Where is the baseline to judge if these incidents are significant? Sample size represents 93 ICUs in Australia.
Research implications	Study needs to be done with controls to determine if the results are significant and the following questions can be asked: What was the percentage of incidents that took place with the experienced staff? What areas of inexperience had the most effect on patient incidents? Would a better skill mix or more nurses improve outcomes?

ID, origin, authors (year)	693, USA, Nash, I.S. *et al.* (1999)
Aims	To determine the magnitude and mechanism of the influence of physician specialty on inpatient mortality for acute myocardial infarction (AMI) *Workforce:* Physicians; secondary *Feature:* Specialisation *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective correlation study 2 Include all AMI admissions in Pennsylvania in 1993, with 24% excluded because they represented secondary admissions resulting from a hospital-to-hospital transfer. 3 30,351 admissions 4 N/A 5 Data from Pennsylvania Health Care Cost Containment Council. Charge data were collected by each hospital as well. Severity of illness data were collected on each patient using the Atlas severity grouping system. Attending physician identity and specialty designation were supplied by each hospital.
Results Quantitative results	The influence of physician specialty on inpatient mortality for AMI was studied. In patients <65 years old, the adjusted odds ratio (OR) for mortality with cardiologist care was 0.89 (p = 0.49) relative to generalist care. In patients >65 years old, the adjusted OR was 0.86 (p = 0.10). Caseload was significantly higher among cardiologists and was inversely related to inpatient mortality.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Severity of illness data were collected using the Atlas severity grouping system. Each patient was assigned a disease-specific 'severity' index, called an admission severity group (ASG) of 0, 1, 2, 3, or 4, based on a variety of clinical abstracted variables. Higher scores were assigned to patients with higher anticipated inpatient mortality and scores were constructed to yield standardised probabilities of death. 2 Mortality analysis was split into 2 strata, defined by age <65 years and age >65 years. 3 Yes 4 Completed. 5 Assignment of types of physicians to AMI patients was not random. Sampling included all AMI admissions except exclusion. 6 One state: Pennsylvania
Commentary	Big sample size. Limitation includes reliance on Pennsylvania Health Care Cost Containment Council for data.
Research implications	The study explains the trend toward better outcomes among AMI patients of cardiologists rather generalists.

ID, origin, authors (year)	134, USA,Palefski, S.S. and Stoddard, G.J. (2001)
Aims	To compare catheter-related complication rates in patients who had infusion devices placed by infusion nurses with complication rates in patients who had devices placed by generalist nurses *Workforce:* Generalist nursing staff and infusion nurses *Feature:* Quality of patient care and complication rates *Outcome:* Rate of leakage, phlebitis, infiltration, complications and period of time vascular access device (VAD) inserted
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Prospective evaluation and literature review of existing patient care 2 Not stated. 3 n=639 patients treated by infusion nurses, n=137 patients treated by generalists: 776 VADs – 442 peripheral–short catheters placed as first, 221 as second and 113 as the third VAD 4 No follow-up 5 Data collection form completed by nurses, 2 hospital sites and 1 agency over a consecutive 3-month period from 1998 to 1999.
Results Quantitative results	Complication rates: 36% of all VADs inserted by generalist nurses were removed because of a complication compared with 20% of VADs inserted by infusion nurses (p <0.001). The relative risk (19.7%/35.8%) for this removal rate was 0.55, and the risk difference (35.8%–19.7%) was 16.1%. A significantly lower incidence of leakage occurred with VADs inserted by infusion nurses (6.4% vs. 15.3%, p =0.001). As verified with the Cox regression, patients in the infusion nurse group exhibited one-third of the risk for VAD leakage (relative risk = 0.33, p = 0.001). The incidence of infiltration was significantly lower with VADs inserted by the infusion nurses (7.5% vs. 13.9%, p = 0.028). From the Cox regression, patients whose devices were inserted by infusion nurses were associated with half the risk of infiltration relative to the generalist nurse group (relative risk = 0.42, p=0.005). In the case of peripheral infusion by specialists, there was a significantly lower rate of leakage, phlebitis and infiltration complications and the VAD remained in place significantly longer.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Cox regression analysis used to control for all sources of imbalance 2 Where appropriate, an adjustment for predisposition to phlebitis (patients with prior VAD who had phlebitis), and adjusted comparison of risk 3 Uniform collection of data via data collection form which listed symptoms and clearly noted definitions for infiltration and phlebitis 4 N/A 5 No randomisation of catheter insertion, which led to differences in patient and therapy characteristics 6 No clear data for reasons of privacy; 2 hospital sites and 1 home infusion agency
Commentary	Specialist credentials determined by nurse manager based on manager's own criteria for skill and experience; therefore, no standardisation of measure. No informed consent and no IRB overview, therefore no naming of specific site or locations – dispersal unclear, as is type of hospital and client population (e.g. Medicare, Medicaid patient mix). One limitation of the study was that the VAD insertions were not randomised between the study groups – a Cox regression model was used to statistically correct the imbalance. The generalist nurses inserted VADs primarily in ER settings (42.5%) and other (33.9%), whereas the specialists worked primarily in Med/Surg (83.5%). This would have had an impact on the number of leakage and other complications due to the unsettled nature of inserting in the ER setting, patient stress, the necessity of moving the patient to a ward etc.
Research implications	Further study of the relationship between standardised specialisation and outcomes. Comparison with the more common model of an infusion therapy team.

ID, origin, authors (year)	347, USA, Philbin, E.F. *et al.* (1999)
Aims	To determine whether there are treatment choices and clinical outcomes differences among patients with congestive heart failure (CHF) treated by cardiologists and by non-cardiologists in the community hospital setting *Workforce:* Physicians; Secondary *Feature:* Specialisation *Outcome:* Hosptial length of stay, mortality, hospital re-admission and quality of life
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A prospective cohort study 2 Included all patients assigned CHF as the primary diagnosis disease in the 10 participating hospitals during two 9-month periods. Excluded those patients with CHF as a secondary diagnosis position and excluded those with incomplete records, incomplete follow-up or no designaton of the specialty of the attending and consulting physicians. 3 2,454 patients 4 6 months after hospital discharge 5 The baseline database of the management to Improve Survival in Congestive Heart Failure study was used. The records of eligible patients were audited by trained personnel immediately after hospital discharge. The survivors were followed up for 6 months after hospital discharge, and they were contacted by telephone three times during this period. The occurrence and cause of re-hospitalisation or death were determined by discussions with the patient, the patient's family, the attending physician, and by a review of hospital admission logs, the patient's hospital records and death certificate. VF or death were determined by discussions with the patient, the patient's family, the attending physician, and by a review of hospital admission logs, the patient's hospital records and death certificate.
Results Quantitative results	Patients with CHF were identified and followed up for 6 months after hospital discharge. Patients who were not treated by a cardiologist (group I; n=977) were compared with patients whose attending physician was a cardiologist (group II; n=419) and patients who received consultative care from a cardiologist (group III; n=1058). Outcome measures were hospital length of stay, mortality, hospital re-admission and quality of life. When compared with group I patients, group II patients were more likely to receive the recommended diagnostic tests and treatment strategies. Group II patients had higher hospital charges, but lower CHF re-admission rates and better post-discharge quality-of-life measures. No differences in adjusted mortality rates were observed.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 All clinical outcomes were severity adjusted. To adjust for the influence of baseline case mix differences on each outcome variable, the significant clinical covariables selected by each of the four outcome measures were entered as independent variables into a regression model for that outcome. 2 No 3 Yes 4 Complete follow-up 5 Not random. Assignment of specialists or non-specialists was by hospital. 6 10 acute care community hospitals in New York
Commentary	The choice of process-of-care markers and quality-of-life measures were subject to the discretion and potential bias of the authors, some of whom are specialists in CHF and may hold views that are different from the views of non-cardiologists. Study sample was focused on patients with CHF as the primary reason for hospitalisation and thus excluded those with CHF as a secondary problems.
Research implications	Future studies can investigate whether a more rigorous compliance with published guidelines by non-cardiologists would offer the same benefits as cardiology specialty care. The relationship between physician specialty, process of care, and clinical outcomes requires further study before effective sweeping health manpower recommendations can be made.

ID, origin, authors (year)	335, USA, Philbin, E.F. and Jenkins, P.L. (2000)
Aims	To explore the relationship between cardiac speciality care and short-term heart failure-related outcomes *Workforce:* Cardiologists, internists, family physicians and 'other' physicians, mixed *Feature:* Specialisation *Outcome:* LOS, mortality, re-admission and the composite of in-hospital mortality and re-admission
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective examination of administrative database 2 All patients with primary diagnosis of heart failure (HF) were included, irrespective of the procedures performed or DRG. Patients who died were eliminated from the analyses of re-admissions. 3 236 acute care hospitals; 10,506 cardiology patients, 28,300 internal medicine patients, 4812 family practice patients and 1308 other patients 4 6 months for re-admission 5 Information on all 1995 New York State hospital discharges assigned ICD-9-CM codes indicative of HF in the principal diagnoses position, re-admissions and speciality of attending physician were obtained from the Statewide Planning and Research Cooperative System (SPARCS) database. The period 1 January to June 1995 was used for discharges and the remaining 6 months for follow-up.
Results Quantitative results	*Unadjusted:* Patients of cardiologists had a mean hospital LOS nearly 1 full day less than patients of internists but equivalent to patients of family practitioners. The other group had the longest mean hospital LOS. Mortality and re-admission rates were similar among all four groups. *Adjusted:* LOS was similar for cardiology and internal medicine patients but lower for family practice patients. Adjusted LOS was highest among other patients. The adjusted OR for death, re-admission and the composite measure were equivalent to cardiology for the remaining three groups.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Age, sex, race and medical comorbidities (using Charlson index) 2 Location (urban vs. rural) and teaching status 3 Yes 4 7084 excluded from analyses because they did not contain physician specialty information. 5 No 6 Not stated.

Outcome	Cardiology		Internal medicine		Family practice		Other	
	unadjusted	***adjusted***	***unadjusted***	***adjusted***	***unadjusted***	***adjusted***	***unadjusted***	***adjusted***
LOS**	8.8 ± 10.7	9.4	9.7 ± 15.0*	9.5	8.9 ± 12.9	8.6*	11.5 ± 14.2*	11.7*
Mortality	6.5%	1.00	6.7%	0.94 (0.85–1.03)	7.2%	0.97 (0.84–1.12)	7.0%	1.12 (0.89–1.42)
HF readmission	28.6%	1.00	28.4%	1.01 (0.94–1.09)	26.8%	0.98 (0.88–1.09)	27.6%	0.99 (0.82–1.19)
Composite measure	33.8%	1.00	33.5%	1.00 (0.94–1.07)	32.7%	0.99 (0.90–1.10)	33.5%	1.06 (0.89–1.25)

* $p \leq 0.01$ for comparison with cardiology, ** $p \leq 0.01$ for comparison across all 4 groups

Commentary	The administrative data lacked disease-specific markers for severity, thus prohibiting disease-specific, risk-adjusted comparisons between patients. Thus the adjusted outcomes reported here should be interpreted with caution because they may inadequately account for the differences among the groups. The outcomes were all in-hospital-based events. The segregation of patients into physician speciality groups was based on the specialty of the attending physician. In some cases there was probably incomplete accounting for the influence of all physicians who contacted patients during their hospitalisation, including consultants and those who performed procedures.
Research implications	Better adjustments are needed in any further analyses.

ID, origin, authors (year)	1137, USA, Posner, K.L and Freund, P.R. (1999)
Aims	To investigate the trends in quality of anaesthesia care associated with changing patterns in a university hospital *Workforce:* Junior doctors (residents and fellows) and senior doctors, specialist nurses (Certified Registered Nurse Anaesthetist or CRNA); tertiary care *Feature:* Specialisation; productivity (measured on a monthly basis by dividing the total attending anaesthesia hours (time units) by the sum of clinical days worked by all attending anesthaesiologists) and concurrency (measured as the number of cases an anesthaesiologist supervises during overlapping time periods) *Outcome:* Complications, adverse events and medical errors. Quality of anaesthesia care measured as monthly rates of critical incidents, patient injury, escalation of care, and human errors per 10,000 cases. Operational inefficiencies and changes in team composition were also measured but will not be reported in this abstraction.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort 2 Patients who underwent anaesthesiology at the University of Washington Medical Center from 1992 to 1997 and the anaesthesia teams working at the hospital during those years 3 Range of 12,970 to 14,886 caseloads between 1992 and 1997 inclusive 4 Complete 5 Productivity and concurrency data were gathered from the Department of Anaesthesiology clinical activity database. Data for quality of care were gathered from the Department of Anaesthesiology CQI (continuous quality improvement) Program database.
Results Quantitative results	Over a 6-year period of changing staffing patterns and increasing anaesthesia productivity, most indicators of the quality of anaesthesia care did not appear to decrease. *Productivity over 6 years measured by mean rate/10,000 cases (95% CI)* **Quality indicators** / **Low productivity** / **High productivity** / ***p*-value** Patient injury / 134 / 38 / $p = 0.002$ Critical incident / 36 / 92 / $p = 0.001$ The rates of escalation of care (mean 289/10,000 cases) and human errors (mean 47/10,000 cases) did not exhibit any statistically significant relation with levels of productivity ($p = 0.345$ and $p = 0.320$ respectively). Quality indicators at different concurrency levels followed similar patters. The patient injury rate decreased ($p = 0.001$), critical incident rate increased ($p = 0.002$), and escalation of care and human error rates were not significant ($p = 0.392$ and $p = 0.069$ respectively).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No adjustments were made with patients or medical staff 2 Productivity levels were constructed by rounding monthly productivity levels to the nearest full hour (integer). Concurrency levels were constructed by rounding to the nearest decimal (tenth). 3 Yes 4 Yes 5 No 6 Results apply to the all anaesthesiologist teams and their patients at the University of Washington Medical Center from 1992 to 1997.
Commentary	Limitations include drawing data from a single academic institution with retrospective data collection and reliance on voluntary self-reporting methods of adverse events and outcomes. Quality measures are aggregated and cannot be interpreted to reflect the relative quality of care provided by solo anaesthesiologists vs. attending-resident or attending CRNA care teams.
Research implications	Is the reduction in adverse outcomes the result of quality improvement efforts in response to CQI reports instead of changes in productivity and concurrency? What other factors could be associated with decreasing rates of patient injury? This study took place over a lengthy time period – many other factors could have influenced results. Studies should be done that also assess patient satisfaction directly from the patient's experience.

ID, origin, authors (year)	694, USA, Regueiro, C.R. *et al.* (1998)
Aims	To explore whether supervision of care of severe chronic obstructive pulmonary disease (COPD) patients by pulmonologists is associated with greater costs or better survival compared to generalists *Workforce:* Physician; tertiary *Feature:* Specialisation *Outcome:* Mortality, resource intensity and hospital costs
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Comparative retrospective study 2 Excluded patients who were admitted to surgical, cardiology, or oncology services. Patients were excluded if they had status asthmaticus, if they were pregnant, non-English speaking, non-resident foreign nationals, transferred from another hospital to a non-ICU setting, diagnosed as having AIDS, hospitalised with an expected length of stay <72 hours, or admitted following head trauma. Eligible patients who were discharged or died within 48 hours of study entry were excluded. 3 866 adults with severe COPD from five academic medical centres; 512 had generalists and 354 pulmonologists as their attending physicians. 4 N/A 5 Data from patients enrolled from 1989 to 1994 in the Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT) which was an observational (phase I) and interventional (phase II) study. Patient charts were reviewed for information about diagnosis, comorbid conditions, and resources utilisation. Data on patients' demographic characteristics, preferences to undergo cardiopulmonary resuscitation in the event of cardiac arrest and patient's functional status was collected by interviews.
Results Quantitative results	14% of patients died within 30 days. There were no differences in resource intensity and hospital costs in those treated by pulmonologists or generalists. Patients with pulmonologist as attending physicians did not experience better survival.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 To adjust for non-random assignment of patients to specialty care, a propensity score was developed to estimate each patient's probability of having a pulmonologist as the attending physician. This propensity score was based on a logistic regression model with specialty of the attending physician as the dependent variable. This score allowed adjustment more fully for differences in case mix between pulmonologists and generalists. 2 Patient's acuity of disease is also considered to give an adjusted and fair comparison. 3 Two data collection methods (observational and interventional); within-method uniformity 4 N/A 5 Not random. The assignment of types of physicians depended on the hospital. 6 Five medical centres; location is not stated clearly (probably in three states: Massachusetts, Tennessee, Virginia)
Commentary	SUPPORT was not specifically designed to compare outcomes among different types of physicians. The assignment of costs to an attending physician is fraught with complexity. Some important data were not collected, such as information on co-management of patients by multiple physicians and information about who actually made certain care decisions.
Research implications	There was no evidence of a survival advantage for patients cared for by pulmonologists.

ID, origin, authors (year)	338, UK, Ridsdale, L. (2000)
Aims	To explore the effects of specially trained nurses working in primary care on epilepsy patients *Workforce:* Nurses; primary *Feature:* Specialty training in management of epilepsy *Outcome:* Patient attendance and satisfaction; level of information and advice provided to patients to enhance self-management.
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Literature review 2 *Inclusion:* All papers published from 1992 to 1999 which included the words 'nurses' and 'epilepsy' in the title or in text words. Hand search for unpublished findings and papers in press. *Exclusion:* papers regarding specialist nurses and other conditions analogous to epilepsy. 3 Number of units not clearly reported 4 Estimated number of articles: randomised clinical trial (4); quasi-experimental (1); surveys (6); audit(1) 5 MEDLINE, Psychinfo, EMBASE, Science Citation Index, Cochrane Database, CINAHL; 1992 through 1999 6 I Studies based on well-designed, randomised controlled trials, meta-analyses, or systematic review; II Studies based on well-designed cohort or case control studies; III Studies based on uncontrolled studies or consensus 7 Investigation of differences and bias: Summarisation of studies findings with reference to the impact of nurse specialists on: Advising on the creation of registers of patients with epilepsy with monitoring in primary care, advice, counselling, and liaison on behalf of patients with medical, professional and social agencies.
Results Quantitative results	Where nurses have been trained in epilepsy care, there is level I evidence that it is feasible for them to set up and run clinics in family practice. Where this has occurred, there is level I evidence of patient attendance and satisfaction, and level I evidence that there has been an increase in the information and advice recorded as being provided to patients. Nurse-run clinics may improve the emotional well-being of some patients. Annual monitoring, more frequent monitoring for patients with poorly controlled epilepsy are recommended.
Commentary	This was a limited review due to the limited range of articles. Authors suggest extending the search to include other nurse-run clinics. Most of the articles were of descriptive data.
Research implications	Impact of nurse-run clinics on self-management of epilepsy. Further analysis of specialist nurse management of patients with chronic conditions. Potential for cost savings and improved outcomes.

ID, origin, authors (year)	584, USA, Rudy, E.B. *et al.* (1998)
Aims	To compare the care activities performed by acute care nurse practitioners and physician assistants (ACNP/PAs) in acute care settings and the outcomes of their patients with the care activities and patients' outcomes of resident physicians *Workforce:* Nurse practitioners, physician assistants and residents; acute care *Feature:* Specialisation *Outcome:* In-hospital mortality, occurrence of drug reaction, completeness of admission note and re-admission rate
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Comparative, retrospective and longitudinal design 2 Included acute care nurse practitioners, physician assistants and residents working in the two hospitals, who agreed to participate in the study and completed the daily log diaries. 3 2 hospitals; 187 patients treated by 16 acute care nurse practitioners or physician assistants; 202 patients treated by a matched group of 53 resident physicians 4 4 data collection points in a 14-month period 5 Daily log diaries were used to compare the activities and tasks performed by ACNP/PAs and resident physicians. Data collected for 1 week every 3 months (i.e. 4 collection points in 14 months). Scores on the Acute Physiology and Chronic Health Evaluation (APACHE) III and the Therapeutic Intervention Scoring System (TISS) were used to describe the acuity of the patients. 7 clinical patient outcomes were collected: length of stay, in-hospital mortality, occurrence of a transfusion reaction, occurrence of a drug reaction, complications with an invasive procedure, completeness of the admission note, and re-admission to the ICU within 48 hours or the hospital with the same or related diagnoses within 2 weeks.
Results Quantitative results	Patient outcomes and care activities performed by acute care nurse practitioners, physician assistants and residents were compared. Residents cared for patients who were older and sicker, cared for more patients, worked more hours, took a more active role in patient rounds, and spent more time in lectures and conferences. The ACNP/PAs were more likely than the residents to discuss patients with bedside nurses and to interact with patients' families. They also spent more time in administrative activities. Patient outcomes differences between the groups are not statistically significant.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 APACHE III and TISS were reported to describe patients' characteristics (age and acuity) by type of providers. However, these numbers were not used to adjust the patient outcome to compare the two groups. 2 No 3 6 resident physicians refused to participate in the study because they were too busy while one of 54 participated residents lost the dairy. 2 ACNP/PAs refused to participate. Loss to follow-up: of the 16 ACNP/PAs, 11 participated in all 4 data collection periods, 2 participated in 3 periods, 1 participated in 2 periods and 2 completed only 1 period. Reasons were leaving the position (n=1), stopped seeing patient (n=1), lost or incomplete diaries (n=4) 4 Partially completed (see above). 5 All ACNP/PAs and resident physicians in the two hospitals were invited to participate in the study but not all of them participated (see note 3 above). 6 Two hospitals only (1 in Pennsylvania, 1 in Ohio)
Commentary	Sample size of number of ACNP/PAs was relatively small but the patient sample size was acceptable. The 7 outcome indicators might not be sensitive enough to set out differences between ACNP/PAs and resident physicians. Also, all daily log diaries were self-reported.
Research implications	Future studies should include more sensitive outcome indicators.

ID, origin, authors (year)	249, USA, Silber, J.H. *et al.* (2000)
Aims	To compare outcomes of surgical patients whose anaesthesia care was personally performed or medically directed by anaesthesiologist or generalist *Workforce:* Anaesthesiologists and anaesthetist or nurse anaesthetist *Feature:* Patient and hospital procedures associated with quality of care *Outcome:* Quality of care outcomes, death rate, in-hospital complications and failure to rescue rate
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort analysis 2 Inclusion: Patients 65 or older undergoing general surgical or orthopaedic procedures. 3 n=217,440 patient procedures: 194,430 directed by anaesthesiologist, 23,010 undirected surgeries in 245 hospitals. (14,137 patients not billed for anaesthesiology were classed as undirected. 4 30-day follow-up after admission 5 Pennsylvania Medicare claims records, longitudinal record created for each patients by appending all medical and surgical inpatient and outpatient claims and physician claims during that time interval. Data also included the American Hospital Association Annual Surveys for 1991–1993 and the Pennsylvania Health Care Cost Containment Council Database for 1991–1994.
Results Quantitative results	Adjusted odds ratios for death and failure to rescue were greater when care was not directed by anaesthesiologists (odds ratio for death = 1.08, p <0.04; odds ratio for failure to rescue = 1.10, p <0.01) whereas complications were not increased (odds ratio for complication = 1.00, p <0.7). This corresponds to 2.5 excess deaths/1,000 patients and 6.9 excess failures to rescue (deaths) per 1000 patients with complications. Both 30-day mortality rate and mortality rate after complications (failure to rescue) were lower when anaesthesiologists directed anaesthesia care.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Case mix adjustment of severity of disease and DRG procedure category 2 Adjustment of other provider characteristics 3 Uniform data collection, creation of comprehensive longitudinal record from a number of databases 4 N/A 5 No randomisation 6 245 hospitals in one state
Commentary	Unbilled cases were either supervised by a physician or a staff nurse anaesthetist employed directly by the hospital or may represent undirected anaesthesiology resident cases. The accuracy of the definitions for anaesthesiologist direction is only as reliable as billing data submitted by caregivers. Limitations: outcomes based on retrospective analysis of administrative claims data.
Research implications	Future work is needed to determine whether the mortality differences in this report were caused by differences in the quality of direction among providers, the presence or absence of direction itself or a combination of effects. The next phase of the study will pursue in-depth, large-scale medical chart review which will provide more detailed information.

ID, origin, authors (year)	88, USA, Singh, V., Gress, D.R., Higashinda, R.T. *et al.* (2002)
Aims	To determine whether outcomes for coil embolisation improved with the experience for the practitioner, after adjusting for the perceived risk of treatment *Workforce:* Doctors (neuroradiologists) ; Secondary care *Feature:* Training of workforce *Outcome:* Complications and adverse events (neurological and non-neurological) and length of stay. The study also measured various hospital costs but those results are not reported in the abstraction.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective 2 Patients who suffered from unruptured aneurysms treated by endovascular means. Inclusion criteria: endovascular coil embolisation of an unruptured aneurysm, age 18 years or older at follow-up, no associated arteriovenous malformation, no subarachnoid haemorrhage from a different aneurysm within 6 months before treatment, and no aneurysm treated on a second occasion within 2 months of treatment. A total of 4 patients were excluded. 3 94 patients, 3 physicians 4 Complete 5 Data came from patient medical records and institutional administrative databases from 1990 to 1997 from the University of California San Francisco Medical Center.
Results Quantitative results	The risk of complications with coil embolisation of unruptured aneurysms decreased dramatically with physician experience, even after adjustment for case complexity. Complications at discharge: n (%) initial cases (first 45)/later cases (remaining 49) *p*-value: 11(24)/5(10) $p = 0.07$ Length of stay total: (n) initial cases/later cases *p*-value: 5/3 $p = 0.23$ Adverse events during initial follow up at hospitalisation: No differences existed in types of adverse events neurological and non-neurological. When analysis of only those adverse events that were neurologically or directly related to the procedure were compared, the proportion of adverse events was 15 (33%) in the initial group and 9 (18%) in the later group. Complications occurred in the 53% of the first 5 cases that each of the three physicians treated and in 10% of later cases ($p < 0.001$). After adjustments for all other predictors, the odds of complication were lower with increasing experience (odds ratio 0.69 for every 5 cases treated; 95% CI 0.05, 0.96; $p = 0.03$); the result corresponded to a 30% odds reduction for complication for every 5 cases treated.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjustments were made for risk assessments, age, female sex, and number of aneurysms, aneurismal location, size, and neck–to-dome ratio. 2 Adjustment was also made for Rankin score at admission (reflects morbidity and resource consumption) 3 Yes 4 Yes 5 No 6 Results are only generalisable to the patients at the UCSF Medical Center with coil embolisation of unruptured aneurysms from 1990 to 1997.
Commentary	Advances in coil technology, catheters, wires, and imaging modalities could account for some of the improvement seen later in the study. The study included experienced neurointerventional radiologists who were trained in other endovascular techniques prior to the study period. Results may not be reproducible to those with less experience in endovascular techniques and may be expected to improve more slowly.
Research implications	A cross-sectional study is needed to make results generalisable, as well as with less advanced endovascular physicians. What causes these adverse events? What are the benefits/drawbacks of providing more extensive training for these physicians? Are there any other patient outcomes that are affected?

ID, origin, authors (year)	61, USA, Siqueria, T.M. *et al.* (2002)
Aims	To compare the laparoscopic donor nephrectomy (LDN) results obtained by two different surgical teams. *Workforce:* Surgeons; secondary care *Feature:* Specialisation of the surgical groups, group 1 consisting of a proficient laparoscopic surgeon assisted by an inexperienced laparoscopic surgeon; group 2 consisting of two proficient laparoscopic surgeons *Outcome:* Donors' post-operational complications
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Observational retrospective cohort study 2 The initial 70 sequential LDNs performed during the study period 3 26 consecutive left-sided LDNs were performed by group 1, 44 cases were performed by group 2 4 17 months. In institution 5 N/A; October 1998 to March 2001
Results Quantitative results	A surgical team composed of two proficient laparoscopic surgeons during the early learning curve of LDN may allow safe and efficient development of a laparoscopic live donor renal transplantation programme. *Numbers of laparoscopic donor complications between group 1 and group 2* **Group 1**: 2 major complications: Splenic injury; left adrenal vein clip dislodgement. 2 minor complications: Duodenal ulcer; left lower lung lobe atelectasis **Group 2**: 2 major complications: Laceration in left renal vein branch; ureteral section below ureteropelvic junction. 5 minor complications: Left testicular edema; left testicular pain; left tight numbness; respiratory distress; self-limited atrial fibrillation
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Donor age or body mass index was taken into account; no difference was found between the groups, but gender was not adjusted for. No mention of the donor's comorbidity or the socioeconomic factors. 2 N/A 3 Uniform (only one institute) 4 Complete 5 N/A 6 One institute
Commentary	The need for verbal instruction and the time execution were minimal compared with the constant direction and guidance necessary with a novice laparoscopic surgeon. This might explain the differences in operative time and estimated blood loss between the results of groups. Only one institution was measured, it is lack of generalisability.
Research implications	Need studies that cover wider geographical areas and with broader sample. Need more considerations on case mix adjustment.

ID, origin, authors (year)	938, USA, Solomon, D.H. *et al.* 1997
Aims	To compare outcomes of care provided by generalists with that provided by specialists for patients with musculoskeletal and rheumatic conditions *Workforce:* Physicians; secondary *Feature:* Specialisation *Outcome:* Patient-centered (clinical outcomes or patient satisfaction); resource utilisation (duration of hospitalisation or cost of care); appropriateness of process of care
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 A critical review of studies comparing generalist with specialist care 2 English-language studies only of patients with rheumatic or musculoskeletal conditions or both 3 Not stated. 4 Studies of generalist and specialist care compared with respect to clinical outcomes, resource utilisation, patient satisfaction, or appropriateness of care. Reportedly 17 studies fulfilled inclusion criteria; however, 16 studies only displayed in review tables. Mixed study types: randomised controlled trial (n=1), no. patients = 608; prospective cohort (n=4), no. patients range 282–1633; retrospective cohort (n=8), no. patients range 19–6183 (n=7) and 14,964,900 (n=1); physician survey (n=3), no. surveyed not known. 5 Sources searched: Medline 1966–1996 6 Practitioners: Description of physician training?; random assignation of patients?; similar settings for comparison groups with respect to organisation and physical environment? Patients: Diagnoses described using standard criteria?; diagnoses similar between providers?; similar patients with respect to demographics, case-mix and comorbidity?; adjustment by study authors for differences in the analyses? Outcomes: Validated outcome measures used?; outcome assessors uninvolved to care of patients and blinded to provider assignment?; criteria used to judge appropriateness based on evidence or consensus? Analysis: Power of study adequate to detect meaningful differences? 7 Investigation of differences and bias: critical review only. No combination of primary studies attempted.
Results Quantitative results	Critical review suggests that clinical outcomes for low back pain seem to be similar across different types of providers. Resource utilisation was higher in patients seen by chiropractors and orthopaedists, while satisfaction was highest in patients seen by chiropractors. Authors conclude that for low back pain generalist care seems to be as effective and less expensive – but less satisfying – to patients. Authors conclude that for rheumatoid arthritis, specialists seem to produce better outcomes. For work-related injuries (no. studies = 1), patients seeing chiropractors lost less work time than did those seeing others providers. For studies of osteoarthritis (no. studies = 2), there were differences in process of care between rheumatologists and primary care physicians. For studies of acute arthritis (no. studies = 5), it appears that rheumatologists give more appropriate care and use less resources than do generalists but there was no clear variation in clinical outcomes.

Commentary	Authors recognise that methodological limitations make interpretation difficult. However, more detail of individual studies could be presented in table form for ease of gathering information and primary studies are summarised in the text. Number of reviewers, and their roles, are not known. Four authors involved. Exclusion criteria were not clear, nor were methodological or validity watersheds/minimums – therefore unclear as to the overall validity score/measure applied to included studies. Methodological assessment was scored on: • practitioners – description of training; similarity of settings; random assignment • patients – description of diagnoses; adjustment for differences • outcomes – validated outcomes; unbiased assessment; evidence-based criteria • analysis – power calculation.
Research implications	Provider comparisons are only emergent in research and methodological standards require development and definition and the authors suggest ideas and propose criteria for making such comparisons with respect to: qualification, experience, setting, provider roles. Authors suggest that assignment to provider should be random, case-mix should be controlled for, demographic and social factors should be controlled for. Also outcome and resource utilisation should be determined using standard instruments while, to improve policy making, a broad range of predetermined outcomes should be studied.

ID, origin, authors (year)	360, UK, Soo, L.H., Gray, D., Young, T. *et al.* (1999)
Aims	To determine whether survival from out-of-hospital cardiac arrest is influenced by the on-scene availability of different grades of ambulance personnel and other health professionals *Workforce:* Allied: paramedics, technicians/medical practitioners (MP) and health professionals (HP): nurses, GPs, police, firefighters, ambulance personnel; primary care *Feature:* Specialisation *Outcome:* Mortality measured by survival rates (survival to hospital and survival to hospital discharge)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective observational 2 Patients who had resuscitation attempted by Nottinghamshire Ambulance Service crew from 1991 to 1994 whose arrests were of cardiac aetiology. The following causes of cardiac arrest were excluded: sudden infant deaths, drug overdose, suicide, drowning, hypoxia, exsanguinations, cerebrovascular accident, subarachnoid haemorrhage, trauma, ruptured aortic aneurysm, and pulmonary thromboembolims. From the total of 2094 patients, those whose patient report forms could not be located were also excluded. 3 Total of 1547 resuscitation patients whose arrest was of cardiac aetiology. The number of ambulance crew increased from 22 to 116 (1991 and 1994 respectively). 4 Complete 5 Data was gathered from the Nottingham heart attack register, the four Nottingham A&E departments based at Queens Medical Centre Nottingham, Kings Mill Hospital Mansfield, and the general hospitals at Newark and Bassetlaw; various coronary care units and intensive care units, the records of each ambulance station dispatch and control unit; A&E record sheets, and hospital inpatient case records. Coroner's records and inpatient case records were also examined to identify all those sustaining a cardiac arrest cause (ICD 390–414 and 420–429).
Results Quantitative results	Resuscitation by a paramedic crew from out-of-hospital cardiac arrest caused by cardiac disease resulted in better rates of survival to both hospital admission and discharge from hospital, compared with technician-only crew. *Out-of-hospital cardiac arrest survival to admission and to discharge with crude and adjusted odd ratios*

	Technician crew	**Medic crew**	**Technician + medic backup**	**Technician + MP**	**Technician + HP**	**Medic+ MP**	**Medic +HP**
Number of survivors to admission (%)	36 (6.9)	86 (15.6)	29 (19.9)	11 (19.6)	20 (20.6)	17 (24.3)	22 (20.6)
Crude odds ratio (95% CI)	1.00 (–)	2.49† (1.65–3.74)	3.33† (1.97–5.65)	3.29** (1.97–5.65)	3.49† (1.92–6.34)	4.31† (2.27–8.19)	3.48† (1.95–6.20)
Adjusted odds ratio (95% CI)	1.00 (–)	6.94† (3.92–12.29)	7.16† (3.61–41.22)	4.22† (1.79–9.96)	5.93† (2.93–12.00)	13.82† (5.91–32.30)	12.38† (5.79–26.46)
Number of survivors to discharge (%)	23 (4.4)	32 (5.8)	7 (4.8)	5 (8.9)	7 (7.2)	11 (15.7)	9 (8.4)
Crude odds ratio (95% CI)	1.00 (–)	1.33 (0.77–2.30)	1.09 (0.46–2.58)	2.12 (0.77–5.80)	1.68 (0.70–4.03)	4.03† (1.87–8.66)	1.98 (0.89–4.41)
Adjusted odds ratio (95% CI)	1.00 (–)	3.55** (1.62–7.79)	1.76 (0.59–5.29)	3.24* (1.03–10.20)	2.79 (0.98–7.94)	20.88† (6.72–64.94)	9.11† (3.12–26.61)

* p <0.05; ** p <0.01; † p <0.001
Medic = paramedic; HP = health professional; MP = medical practitioner

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjustments were made for age, Townsend index (a measure of socio-economic status), number of arrest per year, length of experience of leading crew member, presenting rhythm, travel to hospital interval, at-scene interval, location of arrest, witnessed arrest by bystander, and bystander CPR were adjusted by logistic regression model. 2 Number of arrest per year was adjusted for the changing proportion of paramedic and technical crews. 3 Yes 4 Yes 5 No 6 Results apply to the aggregate cardiac patients who were resuscitated by the Nottinghamshire Ambulance service from 1991-1994.
Commentary	It was difficult to control for selection bias since there was no randomisation. The study failed to cite any limitations. The study made ample adjustments to control for extraneous variables so results have significant validity.
Research implications	Study needs to be repeated with other resuscitated groups (i.e. those excluded from the study). A cross-sectional study is needed to make the results more generalisable. How can GPs be encouraged to undertake training in advanced life support?

ID, origin, authors (year)	404, USA, Stearly, H.E. (1998)
Aims	To determine adverse outcomes associated with intra-hospital transportation of critically ill patients by a specially trained nursing transport team. *Workforce:* Specialist nurses *Feature:* Specialisation of workforce *Outcome:* Complications and adverse events measured by substantial changes in heart rate, blood pressure, intracranial pressure, and oxygen saturation
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective observational 2 Patients being transferred from ICU to radiology suits in the University of Missouri Hospitals and Clinics. ICU stat nurses had to be certified in basic cardiac life support, advanced cardiac life support, and paediatric advanced life support. 3 273 patients; 219 adults (over 17) and 18 children 4 Complete 5 Data came from stat nurses flow sheets; assumed hospital records as primary source, January-July 1996.
Results Quantitative results	Use of a specially trained ICU transport team can substantially reduce the rate of adverse outcomes generated by the transportation of critically ill patients for specialised radiological procedures. The patients moved by the specially trained transport team had a 15.5% overall complication rate, with 10.2% minor, 2.5% moderate (compensated for with medications), and 2.8% severe complications that did not respond to intervention. No medications or therapies were delayed, and only 2 patients (0.8%) had decompensation that required the examinations to be aborted. Reported national complication rates for intrahospital transportation of patients are as high as 75%; the complications include adverse events such as delayed administration of medications, significant changes in vital signs, dislodgement of artificial airways and IV catheters, and cardiopulmonary arrest. Minor complications included a heart change of ±10 beats per minute, blood pressure change of ±10 mm Hg; moderate and severe changes were heart rate change of ±10–20 or >± 20 beats per minute and blood pressure change of ±10–20 or >±20 mm Hg respectively. Intracranial pressure changes greater than 7 mm Hg in the absence of stimuli in which the pressure does not return to the baseline were classified as severe, and any decrease of oxygen saturation to less than 93% saturation is classified as severe. For the entire group of patients no IV or central catheters were dislodged, no endotracheal tubes were displaced, and no additional injuries were sustained by the actual physical transportation of the patients for their various specialised examinations.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 No adjustments were made 2 – 3 Yes 4 Yes 5 No 6 Results apply to 27% of patients at the University of Missouri Hospitals and Clinics and the nurse transport teams that worked with them during the time period.
Commentary	No thorough comparison of this study with national studies. No control group. No statistical analysis. Data were only collected for a short period of time.
Research implications	This study needs to be repeated with: adjustments, controls, and some statistical analysis to determine if the findings are significant. A longitudinal study is needed to rule out the effects of time.

ID, origin, authors (year)	85, USA, Stromborg-Frank, M., Ward, S., Huges, L. *et al.* (2002)
Aims	To describe and evaluate the differences in nursing care provided by oncology certified and non-certified nurses to home-based patients with cancer *Workforce:* Nurses: specialty and registered nurses; tertiary care *Feature:* Specialisation of workforce *Outcome:* Symptom management (measured by pain and fatigue), incidence of adverse events (mainly infection), and episodic care utilisation (measured by unplanned visits to hospital, emergency room, or physician's office/clinic). The study also reports planned admissions and unplanned home visits, but those results will not be included in this abstraction.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort 2 Nurses who worked at home care agencies with a case mix that included a high percentage of patients with cancer and at least 20–25% Oncology Certified Nurses (OCNs) among the RN staff and where the desirable patient outcome could be measured. Only one agency fitted these criteria. Cancer patients' charts were selected from nurses employed by the agency from early 1997 through early 1998. Only patients whose entire episode of homecare from admission to discharge was within the time period were included. In each chart, the primary nurse had to play a significant role in the care (facilitated admission and initial care plan, saw the patient for a significant percentage of visits, and completed discharge). 3 7 certified (6 OCNs and 1 Certified Wound, Ostomy and Continence Care Nurse) and 13 non-certified nurses; 181 patients 4 Early 1997 to early 1998 5 All data were collected from The Home Care Program located in a mid-western metropolitan area.
Results Quantitative results	The study failed to support the hypothesis that nursing care provided by OCNs differed from that provided by non-certified nurses in respect to pain management at and after admission, fatigue assessment at admission, and unplanned visits to care facilities. Results did differ between groups in fatigue assessment after admission and the adverse event of infection. The low numbers of decubitus ulcers did not permit statistical analysis. *Documentation of symptom management and adverse events(certified n=74, non-certified n=107)* **Variable** | **Cared for by certified nurses n (%)** | **Cared for by non-certified nurses n (%)** | ***p*-value** *Symptom management* Pain assessed at admission | 56 (76) | 81 (76) | Pain assessed after admission | 66 (90) | 100 (93) | *p* >0.05 Fatigue assessed at admission | 71 (97) | 99 (93) | Fatigue assessed after admission | 18 (26) | 8 (7) | *p* <0.05 *Adverse events* Decubitus ulcers | 1 (1) | 6 (6) | Infection | 16 (22) | 9 (8) | *p* <0.05 *Site of unplanned visits to care facilities (certified n=43, non-certified n=422)* **Visit type** | **Cared for by certified nurses n (%)** | **Cared for by non-certified nurses n (%)** Hospital – general | 7 (25) | 4(18) Emergency room | 16 (57) | 13 (59) Physicians office/clinic | 5 (18) | 5 (23) *Total* | 28 (100) | 22 (100) Patients also differed in two ways not hypothesised: the patients of certified nurses had a greater number of infections and fewer documented instances of patient teaching regarding infection compared to patients cared for by non-certified nurses.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patients were controlled for site of cancer and presence of secondary diagnosis, as well as age, gender, race, marital status, living arrangements, or source of payment for home care. These characteristics did not differ between the two groups. 2 Mean age of nurses in each group was similar. 3 Yes 4 Yes 5 Yes 6 Results are generalisable to patients and nurses at the home care agency during the time of the study.
Commentary	Patient use as a primary source of data and retrospective collection does not allow for standardised measurement of variables or for control of extraneous variables. Differences may have been detected if the sample was more homogenous in cancer diagnosis, stage of disease, and illness trajectory. Data only collected at one setting. The agency used was in a urban location which could limit generalisability to nurses in rural areas. Also, the sample of certified nurses in the agency was small and may not have been representative of all certified nurses, this limiting the ability to generalise the findings. The two groups may have been different in acuity and risk of infection.
Research implications	Studies should be done to ascertain reliability of data (e.g. do home care nurses tend to only document care that is required and reimbursable? Are certified nurses more concerned with patient's personal quality of life?) Study should be done with a more vigorous control over patient cancer characteristics as expressed in the limitations. A cross-sectional study needs to be done to produce results that would be able to be generalised to a larger population.

ID, origin, authors (year)	262, Finland, Suominen, P. *et al.* (1997)
Aims	To study outcomes in paediatric cardiac arrest patents in an emergency medical system based on staff physicians in an ECU and compare it to literature discussing systems that use paramedics *Workforce:* Pre-hospital emergency care units (PECUs): urban (Advanced Life Support (ALS) service) and rural (Basic Life Support (BLS) service) *Feature:* Cardiac arrest management, BLS decision PECU/Helsinki Area Emergency Medical Air Services (HEMS), decision to continue/stop life support *Outcome:* Mortality and neurological disability.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort review 2 Inclusion: 100 prehospital cardiac arrest patient records for patients less than16 years of age. Exclusions: None. 3 100 cardiac arrest patients, of whom 50 had resuscitation interventions performed. 4 Neurological status evaluated upon discharge from hospital and again at 1 year if category was not good. 5 Hospital records for the 10 years between January 1985 and December 1994 were collected from the files of the pre-hospital emergency care unit (PECU) in Helsinki and the same data were retrieved from the run sheets of HEMS from 15 September 1992 to 31 December 1994.
Results Quantitative results	50 patients were declared dead on the scene (DOS) without attempted resuscitation and CPR was initiated in 50. Sudden infant death syndrome was the most common cause of arrest in the patients (68%) as well as in those receiving CPR (36%). There was a significant association in patients with favourable neurological outcome, with a median duration of CPR of 16 minutes. Only 8 of 50 patients on whom resuscitation was initiated survived. Although pre-hospital care was provided by physicians, the overall rate of survival was found to be equally poor as reported from systems with paramedics. The only major difference between the two teams is the ability of physicians to refrain from resuscitation on the scene when prognosis is poor.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 N/A 2 N/A 3 Uniformity of collection of hospital records and autopsy reports 4 Follow-up to 1 year 5 No randomisation 6 No dispersal, pre-hospital emergency care unit of central hospital site in Helsinki, run sheets from HEMS
Commentary	Sophisticated pre-hospital care improves survival in adult cardiac arrest patients, but the overall survival rate of paediatric cardiac arrest patients remains low. Early and effective advanced life support (endo-tracheal intubations) is important to restore spontaneous circulation of a normothermic patient on the scene. Provision of this service in rural areas may be difficult to justify, given the prognosis.
Research implications	Interesting comparison of different emergency medical systems and outcomes.

ID, origin, authors (year)	66, UK, Thompson, J.A. *et al.* (2002)
Aims	To audit anatomical outcome and complications relating to primary surgery for rhegmatogenous retinal detachments *Workforce:*Non-specialist consultant ophthalmologists, specialist ophthalmologists (defined as having declared a specific interest in retinal detachment surgery and able to perform it) *Feature:* Success rate for detachments of differing morphology *Outcome:* Success of re-attachment, complication rates, variation in outcomes
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional clinical survey 2 Not specified in this paper. 3 768 patients and 167 consultant ophthalmologists performing first surgery for simple rhegmatogenous retinal detachment. 4 1-month and 3-month follow-up to determine reattachment rates 5 Clinical data collected in a national cross-sectional survey of all consultants who performed retinal detachment surgery in the National Health Service, methods presented in previous publication.
Results Quantitative results	Overall re-attachment rate with a single procedure was 77% (95% CI 73.9–80.2). A significant difference was seen in re-attachment rates between specialists and non-specialists, overall and for specific subgroups of patients. Allowing for case mix, there was a significant difference between specialists and non-specialists for grade 2 detachments of 87% and 70% respectively (p <.0001). The largest difference between specialists and non-specialists was observed for retinal detachments secondary to horseshoe tears, 80% and 68% respectively (p <.003). Over one-third of patients had at least one complication reported at some point during the audit period.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Yes, adjustment not specified, case selection not specified. 2 Break/detachment type adjusted for. 3 Not specified. 4 Follow-up to 1 month, and 3 months postoperatively for 732/768. 5 No, non specialists have option of choosing which detachments to refer. 6 Not specified, national audit.
Commentary	Most non-specialists performed few surgeries and therefore the lower success rate may be related to lack of quantity as they may not have sufficient practice to maintain clinical and surgical skills. Retrospective design may lead to under-reportage of complications. The change from non-specialist to specialist practice has significant resource implications.
Research implications	This study provides standards to enable surgeons to audit surgical outcomes for primary retinal detachment repair and identify common categories of failure. A prospective study of whether this set of standards impacts on referral patterns would be worthwhile.

ID, origin, authors (year)	213, Canada, Tu, J.V., Austin, P.C. and Johnston, K.W. (2001)
Aims	To determine the independent impact of surgeon speciality training (vascular, cardiac, or general surgery) on the 30-day risk-adjusted mortality rate after elective abdominal aortic aneurysm (AAA) surgery *Workforce:* Doctors *Feature:* Specialisation of workforce: cardiac (fellowship training in cardiovascular surgery and performed >5 CABG per year during time of study); vascular (fellowship training in either cardiovascular or vascular surgery and performed <5 CABG per year during time of study), and general surgeons (any remaining that did not fill previous categories) *Outcome:* 30-day mortality rate after elective AAA surgery
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective cohort 2 Patients receiving elective AAA surgery in Ontario between 1 April 1992 and 31 March 1996. Patients were excluded with ruptured aneurysms. All surgeons in Ontario who received and passed their fellowship examinations in vascular or cardiovascular surgery and then linked to the patient cohort. 3 Total of 5878 patient cases; 130 surgeons 4 In-hospital 5 The Ontario Health Insurance Plan (OHIP) physician claims database provided the patients who received elective AAA surgery, then linked non-ruptured patients to the Canadian Institute for Health Information hospital discharge administrative database for information on demographics, transfer status, comorbidities, and in-hospital mortality. Out-of-hospital mortality was obtained from the Ontario Registered Persons Database. Physician information was provided by the Royal College of Surgeons and linked to OHIP data.
Results Quantitative results	Patients who undergo elective AAA repair that is performed by vascular or cardiac surgeons have significantly lower mortality rates than patients who have their aneurysms repaired by general surgeons *Crude and risk-adjusted 30-day mortality rates after elective AAA surgery categorised by type of surgeon* **Speciality** / **No. of surgeons** / **Volume of cases** / **Crude mortality rate %** / **Risk-adjusted mortality rate (95% CI) %*** Cardiac / 14 / 270 / 3.3 / 4.0 (1.4, 6.6) General / 53 / 1193 / 6.5 / 6.2 (5.1, 7.3)** Vascular / 63 / 4415 / 3.6 / 3.5(2.9, 4.1) * Adjusted for age, sex, transfer status, and Charlson comorbidity score ** Significantly higher (p <0.05) than the provincial average mortality rate (4.1%) *Multivariate odds ratios for 30-day mortality after elective AAA surgery categorised by type of surgeon* **Characteristic** / **Regression coefficient** / **Odds ratio** / **95% CIs** / ***p*-value** General / 0.4838 / 1.62 / (1.18-2.23) / 0.0030 Cardiac / –0.0731 / 0.93 / (0.45-1.19) / 0.8423 Vascular / / 1.00 / /
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patients were adjusted for age, sex, transfer status, and Charlson co-morbidity score made from 15 secondary diagnosis fields (comorbidities present at the time of hospital admission as reported by attending physician) in the CIHI database. Hospital characteristics were controlled by four types of hospital: teaching, large (>15 cases/year), medium (7–15 cases/year), and small (1–6 cases/year). 2 Adjustments to the odds ratios were made according to annual surgeon AAA volume, hospital type, and patient characteristics. 3 Yes 4 Yes 5 No 6 This study represents all patients in Ontario who had elective AAA surgery without a rupture between 1 April 1992 and 31 March 1996.

Commentary	Because of the use of administrative data, there was not complete information on the clinical characteristics of patients undergoing surgery. It is possible but not likely that patients of general surgeons were sicker than those of the vascular surgeons. Also surgeon classifications were based on fellowship training. It is very likely that there were some older vascular surgeons who were misclassified as general surgeons since the speciality of vascular surgeons was not recognised by the Royal College until 1981.
Research implications	Do number of patients treated in each category have an effect on mortality rates? What would be the effects of restricting operating room privileges for aneurysm surgery to vascular patients? Does physician specialty have an effect on other patient outcomes such as complications or adverse events?

ID, origin, authors (year)	69, UK, Tytherleigh, M., Wheeler, J., Birks, M., and Farouk, R. (2002)
Aims	To assess morbidity, mortality, and cancer-related outcomes after supervised rectal resection for cancer by surgical specialist registrars (SpRs) *Workforce:* Consultants and surgical specialist registrars *Feature:* Training of workforce *Outcome:* Postoperative morbidity, mortality, complications, and other hazards (need to transfuse and anastomotic leak)
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Retrospective 2 Consecutive patients who underwent elective resection of their rectal cancer at the Royal Berkshire Hospital between January 1995 and December 1999. Patients who underwent urgent or emergency surgery were excluded. 11 patients were excluded from the sample because they underwent resection by an unsupervised SpR. 3 205 patients under the care of 6 consultant surgeons (68 underwent a resection by an SpR; 126 by consultants); 6 SpRs (5 in fourth year of training and one it the third year) 4 3-monthly for the first 2 years, 6-monthly for the subsequent 3 years; routine colonoscopy was performed 18 months and 5 years after surgery. 5 Data are assumed to be gathered from hospital records.
Results Quantitative results	Operative and cancer-related outcomes are not compromised by supervised SpR resections of rectal cancer in selected patients. *Summary of morbidity/mortality rates after surgery* **SpRs** / **Consultants** Wound complication: 5 (7%) / 15 (12%) Urine retention/infection: 5 (7%) / 9 (7%) Cardio-respiratory: 5 (7%) / 11 (9%) Anastomotic leak: 2 (3%) / 6 (7%) Other: 4 (6%) / 11 (9%) 30-day mortality: 2 (3%) / 8 (6%) Expected ratio for morbidity was 0.9 for registrars and 1.0 for consultants (p >0.5) *The hazard or odds ratio for consultants:registrars* **Ratio** / **95% CI** / ***p*-value** Survival: 1.3 / – / 0.31 Post-operative complications: 1.4 / 0.7–2.6 / 0.25 The need to transfuse: 1.3 / 0.7–2.3 / 0.45 Anastomotic leak: 2.1 / 0.4–20.6 / 0.5 Post-operative mortality: 5.3 / 0.73–23.5 / 0.1 The odds ratios for consultants to SpRs were always greater but not significantly so. Local recurrence: no difference (median follow-up of 48 months) p = 0.5 Distant recurrence: consultant:SpR hazard ratio 1.1 Survival: no significant difference (p =0.31) SpRs:consultants hazard ratio = 1.3

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patients were controlled for age, gender, but differences were not significant in the survival analysis. Patients were also controlled for Dukes stage, and type of operation performed and whether they had a defunctioning stoma but no adjustments were made. 2 – 3 Yes 4 Yes 5 No 6 Patients who underwent elective resection of their rectal cancer at the Royal Berkshire Hospital between January 1995 and December 1999
Commentary	Possible selection bias. Study failed to report any limitations.
Research implications	What are the outcomes of patients undergoing elective resection of rectal cancer by SpRs without supervision by consultants? A cross-sectional study needs to be done to confirm the results, preferably with a larger sample size. Study needs to adjust for increased risk of males having a postoperative complication.

ID, origin, authors (year)	635, USA, Wallace, M.B., Kemp, J.A., Meyer, F., *et al.* (1999)
Aims	To evaluate the performance and safety of screening sigmoidoscopic examinations by trained non-physician endoscopists in comparison with board-certified gastroenterologists *Workforce:* Specialist physicians (allied), nurse practitioners and physicians' assistants; secondary care *Feature:* Specialisation of workforce *Outcome:* Effective symptom control measured by depth of sigmoidoscopy and number of polyps and neoplastic polyps detected. The study also looked at cost differences but those results will not be reported in this abstract.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective 2 For patients to be eligible for screening they had to be at least 50 years of age, have no new lower gastrointestinal symptoms, no acute cardiopulmonary disease, negative faecal occult blood tests, and no first-degree relative with colorectal cancer at 55 years of age or younger. The three non-physicans had to undergo specific training (withdrawal of endoscope, and a minimum of 100 examinations under supervision of a physician. Physicians had to have completed 2- or 3-year fellowships, previously performed at least 1000 lower endoscopic examinations at a rate of 300 per year, and had to be board certified. 3 3701 patients, 15 gastroenterologists, 1 nurse practitioner, 2 physician assistants 4 In-hospital 5 All data were collected prospectively on a standardised form. Patients were identified by their primary care provider and were contacted by phone to ascertain if they should be included in the study. Study was conducted at the colorectal cancer screening program of Harvard Vanguard Medical Associates, a staff model HMO.
Results Quantitative results	In comparison with gastroenterologists, trained non-physican endoscopists performed screening flexible sigmoidoscopy with similar accuracy and safety. After adjusting for baseline differences in patient age and sex, non-physicans had a slightly shorter depth of examination, but this did not result in a reduction in the rate of neoplastic polyps, which is the primary purpose of screening. *Depth:* Unadjusted – mean depth, physician/non-physician (*p*-value); 55 ± 9 cm/ 52 ± 10 cm ($p < 0.001$). Adjusted – depth (% of examinations), physician/non-physician (*p*-value), ≥40 cm (94%)/(92%) ($p = 0.07$); ≥ 50 cm (94%)/(73%) ($p < 0.001$) Physicians were no more likely to achieve a depth greater than 40 cm than non-physicians after adjusting for baseline characteristics. *No. of polyps (%):* physician/nonphysician (*p*-value), 321 (23%)/619 (27%) ($p = 0.34$) *No. of neopastic polyps (%):* physician/non-physician (*p*-value), 80 (6%)/180 (8%) ($p = 0.35$)
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 and 2 Adjustments were made controlling for the patients' age, sex, and family history (i.e. a first-degree relative with colorectal cancer older than 55 or a second-degree relative with colorectal cancer, or a family history of polyps) 3 Yes 4 In-hospital 5 No 6 This study represents eligible sigmoidoscopy screening patients over the age of 49 with no pre-existing conditions, or family history of colorectal cancer or polyps and gastroenterologists and non-physician endoscopists who have had similar training with those in the study at similar HMOs.
Commentary	Patients were not randomly assigned to examination by physicians or non-physicians but differences were controlled to some degree by multivariate modelling. Non-physicians and physicians did not examine the same patients so proportion of patients with polyps detected by each type of endoscopist cannot be compared directly. Strengths include large sample size, prospective data collection, and its conduct as a part of an institutional colorectal cancer screening programme.
Research implications	This study should be repeated using several different data sources so that results can be more generalisable. This study should be repeated but to control for patient difference between the two groups; the same patients should be used or patients in the two groups should be matched. Is there a difference between detection rates among nurse practitioners and physician assistants? Understand how training and implementation of non-physician endoscopists and other non-physicians can expand health care and make it more universal.

ID, origin, authors (year)	230, USA, Wheeler, E.C. (2000)
Aims	To determine whether differences exist between patients with total knee replacement on hospital units with or without clinical nurse specialists in terms of selected process and outcome variables *Workforce:* Clinical nurse specialists (CNSs); secondary *Feature:* Additional team member; units with CNSs vs. those without CNSs *Outcome:* Length of stay (LOS, days from date of admission to the date of discharge from the orthopedic unit), total LOS (TLOS, LOS plus LOS in the rehabilitation unit) and complications
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, comparative correlation design 2 Inclusion: patients older than 18 who had undergone total knee replacement (TKR) surgery and had been discharged from the hospital no more than 1 year from the start of data collection 3 2 orthopaedic units with unit-based CNSs (hospital 1 and 2) consisting of 64 patients and 2 orthopedic units without unit-based CNSs consisting of 64 patients 4 In-hospital 5 Data were collected using a retrospective chart review.
Results Quantitative results	LOS: ANCOVA with adjustments showed no significant difference in the LOS between the two groups, but patients with CNSs had significantly shorter TLOS than patients without CNSs. Complications: There were 17 preventable complications on units without CNSs and 6 on units with CNSs; no statistical analysis was performed due to the small numbers.

Outcome	**Mean (SD)**		**ANCOVA**
	CNS	***non-CNS***	***(F)***
LOS	4.5 (7.7)	4.72 (1.78)	0.36
TLOS	4.87 (1.43)	6.84	20.62 (p = 0.001)
	CNS (n)	***non-CNS (n)***	
Respiratory infection/pneumonia	5	4	
Deep vein thrombosis	0	0	
Skin breakdown	0	0	
Foot drop/contracture	1	0	
Surgical wound infection	0	2	
Other (UTI, bleeding, fever, drug overdose etc.)	0	11	
Total	6 (9%)	17 (26%)	

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Adjustments were made for age and type of anaesthetic used during surgery. 2 Structural variables considered were: *Institutional* – number of hospital beds, nurse–patient ratio, number of physical therapists, and type of nursing care delivery *Unit demographics* – number of unit beds, nurse–patient ratio, written nursing care guidelines or standards of care for TKR patients, nurses' average years of experience, and nurses' professional educational background. 3 Although the data were collected in the same way there were differences in documentation of the four hospitals. 4 No patient records were eliminated because of severity of illness 5 Subjects were randomly selected from a computer list of patients who had undergone TKR surgery. 6 Three states in north-eastern USA

Commentary	The CNSs had Master of Science in Nursing degrees, were certified orthopaedic nurses, and had been working as CNSs for >10 years. All units had similar nurse–patient ratios (1:5 to 6 for the day shift, 1:6 to 7 for the evening shift and 1:8 to 10 for the night shift). The chart review was retrospective and only looked at nursing activities that were documented. The researcher was not familiar with any of the hospitals, and extraneous variables that may not have been considered could have affected the dependent variables. One hospital did not require nurses to write notes unless there was a specific variance, and this limited the information conveyed. Another limitation was the use of intact groups as a basis for comparison. Although both groups were found to be similar in gender and severity of illness, there were significant differences in age and type of anesthaesia used.
Research implications	Does the skill mix of the other team members make a difference? Was the CNS providing an additional person or being substituted for another member of staff? Does the experience of the CNSs impact on the outcomes?

ID, origin, authors (year)	151, USA, Wu, A.W. *et al.* (2001)
Aims	To examine the relationship of physician specialty to treatment and outcomes of patients with asthma in managed care plans, compared among generalists, experienced generalists, pulmonologists and allergists. *Workforce:* Physicians *Feature:* Specialisation *Outcome:* Treatment indicators included use of corticosteroid inhalers, use of peak flow meters, allergy evaluation, discussion of triggers, and patient self-management knowledge. Outcome measures included cancelled activities, hospitalisation or emergency department visits, asthma attacks, workdays lost, asthma symptoms, physical and mental health, overall satisfaction with asthma care, and satisfaction with communication with physicians and nurses.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 A retrospective, correlation study 2 3 inclusion criteria: aged 18 years or older as of 1 September 1993; enrolled in the managed care organisations (MCOs) at the time of sampling; and at least 2 medical care encounters (ED visits or hospitalisations) with a diagnosis of asthma in the previous 24 months. 3 1954 patients with 1078 matched physicians 4 N/A 5 Information obtained by mailed, self-administered patient and physician surveys.
Results Quantitative results	The differences of treatment and outcomes of asthma patients cared by generalists and specialists were examined. Significant differences were noted for patients of specialists and experienced generalists compared with those of generalist physicians. Peak flow meter possession was reported by 41.9% of patients of generalists, 51.7% of patients of experienced generalists, and 53.8% of patients of pulmonologists or allergists. Compared with patients of generalists, outcomes were significantly better for patients of allergists with regard to cancelled activities, hospitalisations and emergency department visits for asthma, quality of care ratings, and physical functioning.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Multivariable logistic regression analyses were performed and the researchers controlled for demographics, asthma symptoms, presence of COPD, smoking and passive smoke exposure, and comorbid conditions that increase asthma symptoms. 2 To explore possible mechanisms for speciality-related differences, in a second model, the researchers further adjusted for quality of care indicators including possession of an ICS and peak flow meter, adequacy of information about asthma management, discussion of triggers, and allergy testing. 3 No. Information collected by mailed, self-administered surveys. 4 Completed. 5 No. The sample in this study only included those enrolled in MCOs. Therefore, it was composed of mostly white, well-educated adults insured through plans affiliated with prominent US companies. 6 Probably all across USA. Details not stated.
Commentary	Since the sample in this study was composed of mostly white adults insured through plans affiliated with prominent US companies, the findings can only be generalisable to adults treated in managed health care settings but not the whole population. Also, one of the inclusion criteria was at least 2 medical encounters in 2 years before the study, so asthma symptom scores were more severe in this sample than for all adults with asthma.
Research implications	To pursue goals of accountability and information that can support quality improvement, a range of next steps is proposed by this current research team. These steps include attaining a better understanding of the differences in care provided by sub-specialists, asthma-experienced generalists, and generalists with limited asthma experience. A better understanding of referral and care-seeking practices that lead patients with asthma of similar severity to be treated by physicians with different levels of training and experience. Translating the information learned into new guidelines for training and practice should contribute toward a system that manages health care to ensure the best outcomes possible.

ID, origin, authors (year)	346, Wyatt, J.P., Henry, J. and Beard, D. (1999)
Aims	To compare the survival rate of trauma patients treated by A&E consultants with junior doctors *Workforce:* A&E consultants and junior doctors; trauma *Feature:* Specialisation *Outcome:* Survival rate
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 A 5-year prospective, comparative study 2 Included trauma patients in four Scottish hospitals during February 1992 to December 1996 and excluded children aged less than 13 years and elderly patients with isolated neck of femur fractures or isolated fractures of the pubic rami. 3 Big sample size: 10,968 patients (1208 patients treated by an A&E consultant; 9195 patients treated by junior staff) 4 N/A 5 A 5-year study (February 1992 to December 1996); data collection by the Scottish Trauma Audit Group
Results Quantitative results	The group of patients treated by A&E consultants had a significantly higher survival rate (more excess survivors per 100 patients compared to UK average performance) (p <0.05) than the group treated by junior doctors.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 TRISS methodology (using Revised Trauma Scores and Injury Severity Scores) was used to take account of confounding factors and validly compare the management of different groups of injured patients. 2 No 3 All trauma patients other than those excluded (see Methods point 2) were included in the study. The method of data collection is not clearly stated. 4 Completed. 5 Random 6 Four Scottish hospitals
Commentary	This study had big sample size and the TRISS methodology gave a fair comparison of the two groups. However, since the details of how data collected were not stated, limitations of the study could not be commented on.
Research implications	The data presented support the call for A&E consultants to be increasingly involved in the early management of major trauma.

Table A2.12 Operation

ID, origin, authors (year)	633, USA, Baggs, J.G. *et al.* (1999)
Aims	To investigate the association of collaboration between intensive care unit physicians and nurses and patient outcomes. *Workforce:* Physicians (resident and attending) and nurses, ICU *Feature:* Collaboration *Outcome:* Mortality, ICU re-admission
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective, cross-sectional 2 Patients were 18 years+, had been in ICU for 4+ hours. 3 Three ICUs: 97 attending physicians, 63 resident physicians and 162 staff nurses; 1432 patients 4 N/A 5 When patients were ready for transfer from the ICU to an area of less intensive care, self-reported questionnaires (Collaboration and Satisfaction about Care Decisions – CSACD) were used to assess care providers' reports of collaboration in making the trauma decision during 1994 to 1996. Mortality and re-admission rates were provided by the units.
Results Quantitative results	Nurses' reports of collaboration were associated positively with patient outcomes (mortality and re-admission) in medical ICU (p = 0.037), but not associated in surgical and mixed ICU. For each increase of one point in collaboration, the odds of a negative patient outcome were reduced by 4% (OR = 0.96; 95% CI: 0.926, 0.998). Residents' and attending physicians' reports of collaboration were not significantly associated with patient outcomes.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Severity of illness was controlled for in all regression analyses by using the APACHE III predicted risk of mortality and ICU re-admissions from day of ICU admission. 2 No 3 Yes 4 Participants were not followed up. 5 All nurses and physicians in the three ICUs were invited to participate. 6 One area: New York
Commentary	The average response rate from nurses was 94%, 81% from resident physicians (no responses from mixed ICU) and 70% from attending physicians (no responses from medical ICU). The study was conducted in only one geographic area, which could limit any generalisability to that area. The power in some of the individual analyses may not have been high enough to demonstrate statistically significant relationships that were present.
Research implications	In future research, concentration on units with very sick, complex patients and use of patient outcomes in addition to mortality would maximize the opportunity to assess relationships. Collection of data at both unit and individual levels, although requiring more resources, provides stronger, more complete findings, as variables may be influential at either level. Conducting studies in multiple units would allow discrimination between the effects of collaboration and other variables, such as diagnostic diversity and technological availability. To begin to assess causality, intervention studies will be needed. Any intervention to increase collaboration will have to include all providers from the beginning to optimise its implementation.

ID, origin, authors (year)	108, Canada, Doran, D.I. *et al.* (2002)
Aims	To investigate the relationship of nurse structural variables (hospital experience and education), unit structural variables (job autonomy and role tension), patient structural variables (medical diagnosis, loength of stay, age, gender and education) and patient outcome achievement (Therapeutic self-care, mood disturbance and functional status). *Workforce:* RNs and Registered Practical Nurses (RPNs); tertiary care hospital *Feature:* Autonomy, role tension and communication, co-ordination *Outcome:* Therapeutic self-care, mood disturbance and functional status
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional 2 Patients eligible if data collected 24–48 hours prior to expected discharge. Patients must read or speak English, not be cognitively impaired and be able to give informed consent. 3 372 patients, 254 nurses – from 26 general medical/surgical and cardiac units 4 N/A 5 Structured questionnaires, chart audit. Data collected between 1997 and 1998. Job autonomy, Hackman and Oldham's Job Diagnostic Survey (JDS); role tension, Lyon's (1971) Role Tension Scale; co-ordination and communication, instrument developed by Shortell *et al.* (1991); therapeutic self-care, instrument developed by the authors; mood disturbance, Sutherland *et al.* (1989) Linear Analogue Assessment Scale; functional status, instrument developed by the authors.
Results Quantitative results	*Standardised path coefficients demonstrating the effect of structure and process variables on therapeutic self-care, functional status and mood disturbance*

	Therapeutic self-care		**Mood disturbance**		**Functional status**	
	Direct effect	***Total effect***	***Direct effect***	***Total effect***	***Direct effect***	***Total effect***
Nurse and unit structural variables						
Job autonomy	–	0.07	–	-0.01	–	0.02
Role tension	–	0.02	–	-0.02	–	0.02
Role performance variables						
Nurse communication	0.20	0.15	–	-0.05	–	0.05
Co-ordination of care	-0.13	–0.17	–	0.08	–	–0.08

χ^2 = 23.81, df = 28, p = 0.69; adjusted goodness of fit = 0.97; root mean square residual = 0.05.
Dashes indicate no direct/total effect.

Quality appraisal	
1 Case mix adjustment	1 Stratified according to patient condition
2 Other adjustment	2 Unclear
3 Uniform data collection	3 Yes
4 Participant follow-up	4 Participants were not followed up.
5 Random sampling	5 N/A
6 Geographical dispersal	6 One large centre in Ontario, two sites
Commentary	Patient response rate of 73%. Nurse response rate of 35%. Only direct effect shown for nurse communication and co-ordination with therapeutic self-care.
Research implications	Future research on the direct effects of role performance variables on patient outcomes would be of interest.

ID, origin, authors (year)	1167, USA, Knaus, W.A. *et al.* (1986)
Aims	To examine whether differences in the structure (use of unit and administration of the unit) and process (amount and type of treatment, and interaction and coordination of staff) of intensive care influenced the effectiveness of care, as measured by hospital mortality rates. *Workforce:* ICU staff *Feature:* Co-ordination/collaboration/interaction *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Prospective observational study 2 Included ICUs. Excluded coronary care units, patients under 16 years and those with acute burns. 3 13 hospitals, number of patients in study ranged 159 to 1657 per hospital (only one hospital with greater than 500 patients included). Total patients = 5030. 4 N/A 5 Patient mortality rates were provided by hospitals over a 2–10-month period in 1982 (exception of one hospital – 27 months, from 1979 to 1981). Questionnaire to units' medical/nursing director on nature and practice of ICU (staffing, organisation, policies, procedures, educational affiliation, and extent of the critical care personnel's participation in patient care).
Results Quantitative results	The results are discussed in narrative. One hospital had significantly better mortality rates (41% lower) than expected ($p < 0.0001$), compared with another hospital that had significantly inferior (58% higher) than anticipated ($p < 0.0001$). The hospitals above occurred in the same specific diagnostic category; differences are related to interaction and co-ordination of ICU staff – rather than administrative structure, specialised treatment or hospital teaching status.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Stratified patients by individual risk of death using diagnosis, indication for treatment and APACHE II score. 2 Units classified into: Level I – physician directors/qualified designees in units at all times, high nurse–to-patient ratios, and in-unit teaching and research commitments; Level II – full-/part-time physician directors with qualified designees, and high/intermediate nurse-to-patient ratios; Level III – part-time physician directors, relied on coverage by other in-house physicians, and lower/variable nurse-to-patient ratios. 3 Same data collected from each of the hospitals, possibly in different ways 4 N/A 5 Self-selected after initial invitation 6 Across and within several states in North America
Commentary	Includes tables with characteristics of hospitals and patients, the structure and process of services in ICU, predicted and observed mortality rates. Study was published in 1986; it has an effective methodology, yet may not be representative of ICU care today. Looked only at inpatient mortality, but did a preliminary comparison of patient status 6 months post-discharge in 9 of the 13 hospitals, and found no differences in long-term outcome.
Research implications	This study needs to be replicated to reflect current practice. Investigation into the effects of collaboration in areas outside intensive care and on younger patients is needed.

ID, origin, authors (year)	295, USA, MacPherson, D.S. *et al.* (1994)
Aims	To investigate the effect of care by a full-time internist who is co-managing surgical patients with the effect of internist care via consultation using internal medicine subspecialists *Workforce:* Internist, tertiary care *Feature:* Co-management (collaboration as an indirect effect) *Outcome:* In-hospital mortality and length of stay
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Before/after comparison 2 Patients undergoing cardiothoracic surgery at Minneapolis Veterans Affairs hospital 3 165 patients (86 pre-intervention, 79 post-intervention) 4 Unclear 5 Medical records or hospital computerised databases were used. Pre-intervention (Spring 1989) and post-intervention (Spring 1990).
Results Quantitative results	Significant shortening of postoperative length of stay (18.1 days before and 12.1 days after, p = 0.05) and total length of stay (27.2 days before and 19.7 days after, p = 0.03). In-hospital mortality rate for patients undergoing surgery was 8.1% before the intervention versus 2.5% afterward (p=0.17).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Used Charlson comorbidity index. 2 Unclear 3 Data collection may not have been uniform; medical records or hospital computerised databases were used to collect patient data. 4 There was no participant follow-up in this study. 5 Patients were not randomly selected. 6 Limited to one hospital unit.
Commentary	It is unclear whether both medical records and hospital computerised databases were used for each patient or whether this varied by patient. Study only included one internist at one hospital, limits generalisability. Co-managing internist was aware of study which may have influenced their behaviour. Secular trends may explain findings. Included only patients receiving surgery. Restricted to in-hospital mortality. Are the findings the result of the skilled internist or is the improved effect due to the increased availability of another team member?
Research implications	Future studies need to address a larger sample from different geographical locations, and different patient populations.

ID, origin, authors (year)	478, USA, Mitchell, PH. *et al.* (1989)
Aims	To explore the relationship between structural concepts (centralisation, formalisation, expertise and specialisation), processes (workflow and coordination), clinical outcomes (affective outcomes of patient satisfaction, physical outcomes of mortalityand morbidity), fiscal outcomes (costs) and organisational outcomes (performance, climate and satisfaction). Workforce: RNs and speciality/board certified physicians; secondary care Feature: Collaboration Outcome: Mortality, complications and satisfaction Medical and surgical ICU (10 beds) and coronary care unit (CCU) (7 beds), collectively referred to as 'the unit'. Other organisational features reported: work environment, climate and satisfaction, specialisation, turnover, nursing performance.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional 2 All nurses assigned to the unit or 'on-call' floater pool. Patients had to be in the unit for at least 16 hours. 3 42 (82%) nurses, 23 (85%) physicians and 189 patient admissions (192 patient observations) 4 Unclear 5 Organisational features measured using: Moos Work environment Scale (WES) and Charns Organisational Diagnosis Survey (CODS). Unit organisational processes: CODS. Unit organisational outcomes: Nurse Organisational Climate Description Questionnaire (NOCDQ), Minnesota Satisfaction Questionnaire (MSQ), WES and CODS. Clinical nursing processes: Schwirian's Six-Dimensional Scale of Nursing Performance. Clinical outcome: Patient Indicators of Nursing Care (PINC) and satisfaction measured by Hinshaw and Atwood's (1982) adaptation of Risser's instrument. Data from the surveys and questionnaires listed above were collected from August to October 1986 and February to June 1987.
Results Quantitative results	Agreement that nurse/physician collaboration was high (mean 6.1 ± 0.63 SD on a scale of 1 = strongly disagree to 7 = strongly agree for nurses; 4.4 ± 0.58 SD on a scale of 1 = strongly disagree to 5 = strongly agree for physicians) and the unit functioned effectively in patient care (mean 4.57 ± 0.51 SD (1–5 scale)). Conflict with physicians dealt with by constructive confrontation (mean 4.45 ± 1.22 SD (1–7 scale)) in contrast to smoothing over (mean 3.27 ± 1.25 SD), unilateral action (mean 3.38 ± 1.35 SD), avoiding situations (mean 3.24 ± 1.41 SD), bargaining (mean 3.29 ± 0.93 SD) or forcing the issue (3.88 ± 1.6 SD). Mortality: standardised mortality ratio for demonstration unit sample was 51.2% (17 deaths, 33.2 predicted). Ratio was significantly less than 100% (chi-square 7.905, df 1, 0.001 (p < 0.005). Complications: Mean PINC indicated non-resolution of a disease-related problem present on admission (20.58 ± 5.2 SD, n=169) with mean scores not different from those of the critical care comparison sample (19.34 ± 38.66 SD, n=28). Patient satisfaction: demonstration patient group and their families were generally satisfied with nursing care in all three subscale areas (technical–professional, education, trust) with mean ratings of care received about 4 on a scale of 1–5 (very dissatisfied to very satisfied).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patient status measured by APACHE II, requirement for medical therapeutic intensity measured by TISS 2 Demographic characteristics of the hospital: size, type and purpose. Measured at unit level: specialisation, expertise, formalisation (of procedures), decentralisation (autonomy). (Unclear whether adjustments were made.) 3 Yes 4 There was no participant follow-up in this study. 5 Every nurse, physician, patient meeting specified criteria was included unless refusal to participate. 6 Limited to Overlake Hospital Medical Centre (OHMC).

Commentary	Does not mention collaboration with other staff working in the unit, e.g. respiratory therapists, directors. Estimation that subjects comprised 42% of all admissions in study period (25% short stay, 3% refused consent, 30% not approached due to nurses' request). Not representative for patients having drug overdose and those with short stay. Data collected at different periods in different years to determine stability of instruments over time. Used historical controls for which organisational characteristics were unavailable. Single-case design limits generalisability to all critical care settings. Causal inferences cannot be made at a single site, but point to key variables for measurement in multi-site studies.
Research implications	Multiple organisational indicators of high-quality care need to be measured over a wide sample of settings. A comparable control group must be included in future research.

ID, origin, authors (year)	48, Australia, Mitchell, G., Del Mar, C. and Francis, D. (2002)
Aims	To assess the efficacy of formal liaison of GPs with specialist service providers on patient health outcomes. *Workforce:* General practitioners and specialists (medical and nursing) *Feature:* Formal liaison (any formal arrangement linking GPs with specialist practitioners in the care of the patient, e.g. conferences, shared consultations, formal shared arrangements) *Outcome:* Physical (asthma symptoms, control of hypertension, creatinine and HbA1 levels in diabetics), functional (activities and use of services by schizophrenics), re-admission rates, patient satisfaction and referrals.
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 Inclusion: Controlled or randomised controlled trials involving close co-operation between specialists and GPs. 3 Total participants; intervention (963), control (899). 4 7 experimental studies: RCT (5), clustered randomisation (1) & pragmatic controlled trial (1). 5 Sources searched: MEDLINE (1966–2001), EMBASE (1980-2001), CINAHL (1982-2001), PsychINFO (1984-2001) & Cochrane Library (database of systematic reviews and controlled trials register) –August 2001. Search strategy based on the EPOC group from the Cochrane Collaboration. Hand searching of reference lists. 6 Strategy used by Australian National Health and Medical Research Council. Attention to: recruitment strategy, randomisation procedure, presence and method of blinding, loss to follow-up & method of analysis (intention to treat or not) 7 Investigation of differences & bias: Findings from the 7 studies were combined using narrative alone. All of the studies differed in their patient group, the populations investigated in each primary study were: frail aged, routine orthopaedics, asthmatics, hypertension, diabetics, chronic schizophrenics, and chronic mentally ill.
Results Quantitative results	No consistent benefit with chronic or complex cases (found in all 4 studies that measured physical symptoms) No effect on functional outcomes in chronic psychiatrically ill patients (found in the 1 study that measured this outcome) Greater patient satisfaction when collaboration is present (found in all 4 studies that measured this outcome) No improvement in re-admission rates (found in both of the 2 studies that measured this outcome)
Commentary	Two independent reviewers Differences in illness groups investigated by the studies, diversity of settings and analysis prevented statistical pooling. Small number of studies limits the strength/weight of the findings. Lack of double blinding possibly resulting in reporting bias. Characteristics of practitioners involved in intervention group could affect outcome. All studies were reported to be of adequate quality; analysis based on intention to treat in 5/7 trials, baseline characteristics noted in all studies. Origins of the studies are not explicitly stated.
Research implications	Need for greater volume of well-designed studies in this area.

ID, origin, authors (year)	422, USA, Young e*t al.* (1997)
Aims	To investigate whether low outliers and high outliers (for hospital mortality/morbidity) distinguished by the number and variety of co-ordination practices they use influences patient outcomes *Workforce:* Surgical staff; secondary care *Feature:* Co-ordination of work responsibilities *Outcome:* Mortality and morbidity (occurring in 30 days following index operation) Experiences of the National Veterans Affairs Surgical Risk Study (SVASRS) to highlight best practices in the co-ordination of surgical care.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional 2 Vreterans Affairs surgical services. All patients received major surgical procedure. 3 Site visits to 20 surgical services (mortality/morbidity rates significantly higher than expected – high outliers (10); mortality/morbidity rates significantly lower than expected – low outliers (10)). Clinical and outcome data collected prospectively from 87,000 patients. 4 Site visits were conducted 12 months after patient outcome data had been collected. 5 In-depth on-site assessment of surgical services. 2 days per assessment. Clinical and outcome data from 44 surgical services in SVASRS from October 1991 to December 1993.
Results Quantitative results	The results were presented as narrative alone. Low outliers used a greater number and a greater variety of co-ordination practices for each of the three work activities studied (general administration, direct patient care and graduate medical education). Effective co-ordination practices were evident in high outliers, yet not as frequently as in low outliers. Low outliers: high level of interaction among different types of surgical staff at administrative and patient care levels, opportunities for one-to-one discussions and group meeting among staff, mechanisms in place for training staff and standardising work processes. High outliers: communication and collaboration among surgical staff were weak, few structures or processes were in place to manage the interdependencies that exist among different types of surgical staff, poor opportunities for staff training and standardised work processes.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Unclear 2 No 3 Site visits were carried out with a uniform set of interview protocols, clinical and outcome data from 44 surgical services in SVASRS. 4 Participants were not followed up. 5 Site visit team members and the staff at the 20 participating surgical services were blinded to the outlier status of the services during the site visits. 6 Unclear
Commentary	Study includes description of how expected mortality/morbidity rates are calculated. To ensure that site visit data reflected surgical services at the time the outcome data were collected, any relevant changes at the surgical services since the beginning of the data collection phase for patient outcomes were noted. Small sample size and all participants being members of one hospital system limit generalisability of findings. The study details examples of best co-ordination approach under the three work activities investigated; examples are given for standardisation of work, standardisation of skills, supervision and peer interaction. Mortality/morbidity were only measured in the short term (30 days post-operation); long-term effects of staff co-ordination not investigated.
Research implications	Investigation of co-ordination of care across multiple treatment levels and among health care professionals with very different clinical backgrounds and expertise is required. Research is needed in different settings, with larger samples and longer outcome data collection periods to allow for greater generalisability of findings.

ID, origin, authors (year)	240, South Africa, Zwarenstein, M. and Bryant, W. (2000)
Aims	To assess the effects of interventions designed to improve nurse-doctor collaboration. *Workforce:* Doctor, nurse *Feature:* Collaboration ('to work jointly' – sharing of information, co-ordination of work, joint decision making) *Outcome:* Length of stay, number of visits, unplanned re-admission, satisfaction, accidents and complications, mortality. Other outcomes investigated but not of relevance to this scoping study: adherence to treatment guidelines; resources use; changes in: communication, sharing, power dynamics, mutual respect, uptake of effective therapies
Methods 1 Design 2 In/exclusion criteria 3 Number of units 4 Individual study design 5 Sources searched 6 Validity criteria for primary studies 7 Method of combining primary studies	1 Systematic review 2 Inclusion: RCTs, controlled before-and-after studies, and interrupted time series if validity ensured by EPOC. Nurses and doctors sharing the care of patients (primary or hospital care setting), exclusively, or in multidisciplinary team. Explicit aim of primary studies: collaboration between nursing and medical profession. Collaboration interventions: training workshops, reorganisation of wards into smaller teams, meetings. Must include one or more of the outcomes listed above. Exclusion: substitution, specialised teams/units. 3 Total participants = 1945; US study 1102; Thai study 843 4 2 experimental studies: RCT (US)(1) & controlled before-and-after study (Thai)(1) 5 Sources searched: The Cochrane Library (CDSR, CCTR and DARE), EPOC register (& register of studies awaiting assessment), MEDLINE (completed in November 1999). Used Cochrane search strategy for controlled trials & MeSH heading inter-professional relations & free text terms. 6 Experimental studies for which validity must be ensured by EPOC. 7 Investigation of differences & bias: The studies were not combined, they were discussed separately
Results Quantitative results	Shortened length of hospital stay was found in one of the 2 primary studies (US)(reduced from 6.06 to 5.46 days), the other study (Thai) found a shortened length of stay in the intervention group when excluding in-patient deaths (intervention ward 10.5 days, control ward 11.9 days). No statistical differences in mortality rates were found in both of the primary studies. No studies were identified that measured the other patient outcomes stated above.
Research implications	Studies independently reviewed by two reviewers for first 20 studies, then individually assessed. Only two studies included in the review, from very different economically developed countries.
Commentary	There is a very limited amount of research on collaboration under the stated study designs. It is important that further research is conducted in this field in all countries.

Table A2.13 Well-being

ID, origin, authors (year)	738, USA, Dugan, J. *et al.* (1996)
Aims	To investigate the relationship between increased levels of stress and burnout and increased nurse injuries, patient incidents, personal incidents and staff turnover *Workforce:* Staff nurses (full-/part-time RNs and licensed practical nurses (LPNs)); secondary care *Feature:* Stress, burnout *Outcome:* Medication errors and patient falls
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Cross-sectional 2 Staff nurses working on units with 8+ nurses. Within a given month a minimum of 5 surveys had to be returned from an individual unit for their inclusion in the analysis. 3 19 hospital units; total of 601 surveys completed 4 Unclear 5 Postal survey (symptom-based stress survey score, including self-reporting Stress Continuum Scale (SCS)) distributed to nurses at 3 consecutive monthly intervals. Patient data from hospital departments (nursing services and risk management).
Results Quantitative results	*SCS and stress survey scores correlated with patient incidents (n=48)* **SCS** (n=48) / **Month 1** (n=19) / **Month 2** (n=16) / **Month 3** (n=13) / **Stress survey** (n=48) Medication errors: 0.40* / 0.52* / 0.29 / 0.42 / 0.23** Patient falls: 0.33* / 0.41** / 0.36 / 0.21 / 0.07 * significant at 95% CI ** significant at 85% CI Relationship between SCS scores and patient incidents were consistent over time. The linear increase in patient incidents (related to increased SCS scores) was significant at the $p = 0.02$ level ($F = 6.08$, $df = 1.41$).
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No 2 No 3 Yes 4 Return rate of questionnaires from first month 293 / 600 (49%); second month 32%; third month 26%. 5 All nurses eligible for inclusion were provided with a survey. 6 Unclear
Commentary	Small fluctuations between monthly correlations attributed to small base sizes. Poor rate of follow-up attributed to primary vacation period over summer months. Differences in patient case mix may have affected occurrence of patient falls. Dependent variables calculated as percentage of occurrences per number of nurse shifts, allowing for comparison of incidents among all hospital units regardless of staff size.
Research implications	A larger volume of research is required investigating the effects of staff stress on patient outcomes; the inclusion of patient satisfaction would be interesting. Adjustment for patient case mix needs to be considered in future work.

ID, origin, authors (year)	445, USA, Goodell, T.T. and Coeling, H.V.E. (1994)
Aims	To explore the relationship between quality of care and nurses' job satisfaction and the relationship between patient satisfaction with nursing care and nurses' job satisfaction *Workforce:* Nurses (RNs and licensed practical nurses); secondary care *Feature:* Job satisfaction *Outcome:* Patient satisfaction
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Cross-sectional, pilot study 2 Unclear 3 33 nurses, 168 patients 4 N/A 5 Patient satisfaction measured with Patient Satisfaction Instrument (PSI) developed by Risser (satisfaction with nurses' technical, educational, and interpersonal skills). Nurse job satisfaction measured using the Index of Work Satisfaction (IWS) developed by Slavitt *et al.*
Results Quantitative results	*PSI subscale correlation coefficients** **IWS Subscales** / **Technical** / **Educational** / **Personal** Pay / –0.0976 / –0.0313 / –0.0884 Professional status / –0.0385 / –0.0666 / –0.1543 Interaction / 0.1969 / 0.0519 / 0.0692 Task requirements / –0.2192 / –0.2656 / –0.2948 Organisational policies / 0.0765 / –0.0084 / –0.1383 Autonomy / –0.0393 / –0.03 / –0.1057 *in all cases, $p > 0.05$
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Unclear 2 Unclear 3 Yes 4 There was no participant follow-up in this study. 5 Stratified random sample 6 Urban mid-western US teaching hospital
Commentary	To eliminate bias against illiterate or disabled people, patients given the option of self-administering the questionnaire or being interviewed by a trained nurse investigator. No control for number of days cared for by same nurse. Limited detail concerning this pilot study was presented in this paper.
Research implications	Future research needs to include more than one centre in its sample to increase the generalisability of its findings.

ID, origin, authors (year)	400, Canada, Leiter, M.P. *et al.* (1998)
Aims	To examine the relationships of nurse burnout, intention to quit, and meaningfulness of work with patient satisfaction with nursing care, physician care, information provided and co-ordination of care and outcomes of the hospital stay assessed post-discharge *Workforce:* Nurses; tertiary care hospital *Feature:* Nurse burnout, intention to quit and meaningfulness of work *Outcome:* Patient satisfaction with: nursing and physician care, information provided, co-ordination of care, and outcomes of hospital stay
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Cross-sectional 2 Unit included in analysis if at least 3 patient satisfaction survey responses. 3 16 inpatient units from 2 hospital sites; 605 patients and 711 nurses 4 Unclear 5 Nurses questioned using the Maslach Burnout Inventory–General Survey (MBI–GS), Conditions for Self-Management Scale. Patients surveyed using the Patient Judgements of Hospital Quality questionnaire post-discharge.
Results Quantitative results	*Correlations among nurses' scores and patient satisfaction ratings* Number of units = 16. **p* < 0.05; ***p* < 0.01. Presents scatter plot to show correlation between nurse exhaustion and patients' ratings of overall nurse satisfaction.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No 2 No 3 Yes 4 Unclear 5 N/A 6 Limited to two sites in central Canada
Commentary	Loss of potential data through poor correspondence between two surveys conducted independently. Restricted range of values for patient satisfaction measures, possible biased responses. The study does not appear to have made any adjustments.
Research implications	Need for integration of staff survey with patient satisfaction survey.

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1 Nurse exhaustion	–	0.80**	–0.23	0.67**	–0.77**	–0.73**	–0.72**	–0.55**	–0.70**
2 Nurse cynicism		–	–0.24	0.54*	–0.71**	–0.53*	–0.43	–0.30	–0.46
3 Nurse professional efficacy			–	–0.62**	0.52*	0.06	0.26	0.26	–0.09
4 Nurse intension to quit				–	–0.52*	–0.53*	–0.65**	–0.58*	–0.62**
5 Nurse work meaningfulness					–	0.79**	0.65**	0.53*	0.67**
6 Rating of nurses						–	0.70**	0.65**	0.89**
7 Rating of doctors							–	0.79**	0.77**
8 Rating of information								–	0.69**
9 Rating of outcome									–

ID, origin, authors (year)	116, USA, Tzeng, H.M. *et al.* (2002)
Aims	To investigate the relationship among staff nurses' assessment of organisational culture, job satisfaction, inpatient satisfaction with information about home care and follow-up and general inpatient satisfaction with nursing care *Workforce:* Staff nurses, tertiary centre *Feature:* Job satisfaction *Outcome:* In-patient satisfaction with information about home care and follow-up. Also looks at causal relationship: (a) perception of effective organisational structure = high job satisfaction; (b) high job satisfaction = high patient satisfaction; (c) high satisfaction = high general patient satisfaction. This abstraction is concerned with (b).
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Exploratory study/cross-sectional 2 Registered nurses performing direct patient care appointed to the unit for at least 6 months prior to data collection. Temporary or 'floating' staff were excluded. Patients who had been hospitalised for at least one night (aged 17 years+ at time of study). 3 17 units: adult medical/surgical (13), adult psychiatric (2) and gynaecology/obstetric (2). 520 nurses and 345 patients were included in the sample. 4 4–6 weeks 5 Secondary data from large ongoing study (Redman and Ketefian, 1995). Nurse Assessment Survey (NAS) scales. Nursing Services Inpatient Satisfaction Survey (NSISS) and demographic sheet to discharged patients.
Results Quantitative results	Job satisfaction and satisfaction with home care and follow-up; correlation coefficient = 0.60 ($p < 0.01$) Job satisfaction and general inpatient satisfaction with nursing care; correlation coefficient = 0.21 ($p > 0.05$) Direct effect of inpatient satisfaction with home care, with job satisfaction: $\beta = 0.597$
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Unclear 2 Unclear 3 Yes 4 Follow-up letter sent 4-6 weeks after initial mailing 5 All eligible participants were included in the study 6 One centre in the mid-west USA
Commentary	Response rates reported are poor, 28% for nurses and 36% for patients. Job satisfaction was reported as predicting patient satisfaction well and positively. Patient perceptions and expectations may vary according to socio-demographic and clinical characteristics. Patient responses often skewed towards the extreme categories (highly satisfied/most dissatisfied). Patient responses may be different pre-/post-discharge.
Research implications	Future research must attempt to gain a greater response rate from its participants. The causal relationship between staff job satisfaction and patient outcome must be investigated, ensuring that there are no confounders influencing the findings.

ID, origin, authors (year)	323, USA, Weisman, C.S. and Nathanson, C.A. (1985)
Aims	To examine the relationship between job satisfaction and client satisfaction *Workforce:* Nursing staff (RNs); tertiary care *Feature:* Job satisfaction (relationships with co-workers and patients, work content, supervision, and resources available in the job) *Outcome:* Client satisfaction (client feels that their goals in attending the family planning clinic have been met) Also looks at rate of client compliance.
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Longitudinal observational study 2 Unmarried women under the age of 20 making their first visit for contraception during a 10-month period in 1980–1981. 3 77 county health departments, 344 family planning and community health nurses (RNs), 2,900 clients (baseline interviews). 4 Data on subsequent contraceptive use obtained at 6 and 12 months. 5 Baseline interviews with clients, follow-up data from telephone interviews. Surveys with staff.
Results Quantitative results	Higher clinic staff job satisfaction levels predict higher client satisfaction levels ($p < 0.001$). The regression coefficient ($\beta = +0.32$) is considerably larger than the correlation coefficient ($r = +0.24$). Staff job satisfaction is the strongest predictor of client satisfaction in the equation.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 Adjustments were made for: nurses' mean age, percentage of nursing staff with teenage children, hierarchical level (degree of autonomy exercised by the clinic), staff conflict and nursing influence; by investigating any correlation with the patient outcome. 2 As above 3 Yes 4 Unclear 5 N/A 6 21 of 23 counties in Maryland
Commentary	Published in 1985, now18 years old and perhaps no longer generalisable. Response rate of 86% of all nurses who worked in family planning clinics during the study period. Other organisational variables could have been measured. The observed relationship between staff satisfaction and client satisfaction levels may be spurious, as a variable was not measured that could account for the association. Staff satisfaction was presumed to reflect the climate in which provider–client interactions take place; however, client satisfaction may influence the level of staff satisfaction – both factors could be mutually reinforcing. Focused on between-clinic differences
Research implications	Need to look at within-clinic differences in job satisfaction.

Table A2.14 Human resources and policy issues

ID, origin, authors (year)	1127, Canada, Bell, C.M. and Redelmeier, D.A. (2001)
Aims	The authors hypothesised that there would be no difference in aggregate mortality between patients admitted at weekends and those admitted on weekdays, but there would be between three specified conditions that were theorised to accentuate the consequences of lower staffing. *Workforce:* All staff, acute care hospitals *Feature:* Human resources (HR); weekends vs. weekdays (weekend = period from midnight on Friday to midnight on Sunday. For patients transferred between hospitals, the day of admission = initial day presented) *Outcome:* Mortality (prespecified = ruptured abdominal aortic aneurysm, acute epiglottis and pulmonary embolism; control = acute myocardial infarction, acute intracerebral haemorrhage and acute hip fracture; most frequent causes of death: every ICD-9 (The International Classification of Diseases, 9th Revision, Clinical Modification) diagnosis was ranked according to the total number of in-hospital deaths and death 2 days after admission and from this list selected the 100 diagnoses that caused most deaths).
Methods	
1 Design	1 Non-experimental, retrospective
2 In-/exclusion	2 All mortality regardless of whether the patient had died in hospital, had been discharged home, or had been transferred to another facility. Elective admissions, urgent referrals, elective transfers and births were excluded.
3 Sample size	3 3,789,917 hospital admissions
4 Follow-up time	4 In-hospital and 2 days after discharge
5 Data collection: source and period	5 Hospital discharge data were obtained from the Canadian Institute for Health Information for the period 1 April 1988 to 31 March 1997.

Results
Quantitative results

For the prespecified conditions the mortality rate among patients admitted on a weekend was higher than that among patients admitted on a weekday.

Condition	No. of admissions	Mortality rate		Odds ratio (95% CI)		Death 2 days after admission (or 95% CI)
		Weekday	*Weekend*	*Unadjusted*	*Adjusted*	
Ruptured abdominal aortic aneurysm	5,454	36	42 (p <0.001)	1.32	1.28 (1.13–1.46)	1.35 (1.15–1.52)
Acute epiglottitis	1,139	0.3	1.7 (p = 0.04)	5.47	5.28 (1.01–27.50)	10.47 (1.21–90.65)
Pulmonary embolism	11,686	11	13 (p = 0.009)	1.25	1.19 (1.03–1.36)	1.39 (1.14–1.69)
Myocardial infarction	160,220	15	15	1.02	1.03 (1.00–1.06)	–
Intracerebral hemorrhage	10,987	44	44	1.01	1.01 (0.93–1.11)	–
Acute hip fracture	59,670	7	6	0.95	0.97 (0.90–1.04)	–

For 23 (cancer of the trachea, bronchus, or lung; secondary cancer of the respiratory or digestive tract; chronic ischemic heart disease; cardiac dysrhythmia; unspecified condition requiring aftercare; colon cancer; secondary cancer at other specified sites; aortic aneurysm; pancreatic cancer; breast cancer in women; general cardiovascular symptoms; prostate cancer; stomach cancer;cancer of the rectosigmoid or anus; acute pulmonary heart disease; cancer of the brain; cancer of the lvier or intrahepatic bile ducts; renal failure; myeloma or immunoproliferative cancer; intracranial hemorrhage; intestinal hemorrhage; intestinal disorder; cardiac-conduction disorder; leukemia) of the 100 most frequent causes of death admission at weekends was associated with a significant increase in mortality.
Conversely weekend admission was not associated with a significantly reduced mortality rate for any of the conditions.
Analyses of deaths within two days after admission: All possible diagnoses were included and there was a small increase in mortality among patients admitted at weekends (1.8% vs. 1.6%, p <0.001); when only the 100 most frequent causes of death were included 26 conditions were associated with a significant increase in mortality with weekend admission, and no condition was associated with a significant decrease in mortality at weekends.

Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	1 Patient characteristics: age and sex; comorbidity – Charlson comorbidity index (weighted index of the number of serious coexisting disease on a scale of 0 to 8) 2 None 3 Yes 4 Not stated 5 No 6 Ontario
Commentary	The severity of the patients' illness was not considered. Administrative data were used which may have included coding errors. The analysis did not account for statutory holidays. The study excluded deaths declared by paramedics outside the hospital, which may be more common on a weekend and therefore underestimated the differences in mortality. The authors only focus upon in-hospital mortality and do not allow for consideration of the timeliness of care, patients' degree of satisfaction and other aspects of the quality of medical care. For 77 of the 100 conditions that accounted for the largest numbers of inpatient deaths, admission on a weekend was not associated with a significantly higher rate of death than was admission on a weekend. Administrative data cannot account adequately for differences in the severity of illnesses. Of the 23 conditions for which an association between weekend and mortality was found, more than half were cancers. The authors suggest that reduced staffing on weekends may be to blame, but provide no data to investigate this claim.
Research implications	Are patients who are admitted on weekends sicker than those admitted on weekdays? Do fewer people work in hospitals at weekends? Do those who work at weekends have less seniority and experience than those on weekdays? Does the workforce at weekends provide cover for other health professionals and is it consequently less familiar with the patients? Are there fewer supervisors at weekends and do they have to oversee the work of others they do not know well? Could unmeasured differences in the severity of disease explain the association between weekend admission and increased mortality? Could problems with the quality of care result in higher mortality among patients admitted from the emergency department at weekends than among those admitted on weekdays? If so, how? What was the mix of the staff on the ward at weekends vs. weekdays? Experience, training and education preparation?

ID, origin, authors (year)	482, USA, Halbur, B.T. and Fears, N. (1986)
Aims	To simultaneously investigate the impact of nursing home and aggregate resident characteristics and that of nursing personnel turnover rates on resident discharge and mortality *Workforce:* Nursing, nursing homes *Feature:* Human resources; turnover (the proportion of nurses and aides who voluntarily terminate their employment) *Outcome:* Mortality
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	 1 Non-experimental, secondary analyses of data 2 Not stated 3 122 nursing homes 4 In-hospital 5 A 1978 study of annual turnover rates for nursing personnel and a 1979 follow-up study of resident outcomes provided data for the analysis.
Results Quantitative results	Nurse aides, who provide most direct resident care, contributed disproportionately (double that for RNs) to turnover among nursing personnel. Their average rate of turnover was about double that for registered nurses (RNs). Only nursing home characteristics were important for understanding residents' death rates. **Variable** — **Correlation coefficient (mortality)** Log registered nurse turnover rate — 0.12 (NS) Log licensed practical nurse turnover rate — –0.09 (NS) Log nurse aide turnover rate — –0.02 (NS) As turnover was not found to be correlated to mortality it was not included in the regression analyses.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 No 2 No 3 Unsure 4 65% response rate was achieved but no other information is given. 5 No 6 State-wide - North Carolina
Commentary	Homes that had high turnover rates for one type of nursing personnel had higher rates of turnover for other types. This study showed that turnover rates, at least for RNs, are positively related to resident discharge rates. A 1978 study provided data on turnover rates and a 1979 study provided the resident outcomes data and it is unclear if the 2 studies are measuring the outcomes over the same period.
Research implications	Further research is needed to re-examine the relationship among nursing personnel turnover rates and resident outcomes and in so doing select a large, nationally representative sample of nursing homes, use refined outcome measures, and examine other nursing home characteristics. It would be interesting to repeat this study in a hospital and explore if the same relationships hold in that setting. This study looked at mortality as a dependent variable and investigated various factors that affect this outcome, so further research is needed to look at the relationship between turnover and patient outcomes with adjustment for age, severity etc. Are there any differences between the staff that come in and out?

ID, origin, authors (year)	561, USA, Weinburg, A.D., Lesesne, A.J., Richards, C.L. *et al.* (2002)
Aims	To determine whether admissions to a subacute unit received equivalent care on weekdays as opposed to on weekends with regard to certified nursing assistant and licensed nurse staffing levels *Workforce:* Certified nursing assistants (CNA), licensed practical nurses (LPN) and Registered Nurses (RN), university-affiliated nursing facility *Feature:* Time and day of admission; weekends (Saturdays and Sundays) vs. weekdays (Tuesdays and Thursdays) *Outcome:* Medication errors and falls (presence of required daily nursing note and documentation of meals eaten were also reported but will not be included in this abstraction).
Methods 1 Design 2 In-/exclusion 3 Sample size 4 Follow-up time 5 Data collection: source and period	1 Non-experimental, prospective observational 2 All admitted patients who spent at least one week in a subacute unit during the study period 3 31 residents 4 In-hospital 5 One of the primary authors obtained information by prospectively reviewing charts and medication administration records at least three times per week. A standard data form was used to collect information and staffing levels were ascertained by direct observation or by telephone inquiry during the day shift on the reviewed days. Data collection: January and July 2000.
Results Quantitative results	A total of 60% of weekends had decreased numbers of LPNs assigned to work in the subacute unit. Overall, the number of LPNs was greater on weekdays than at weekends (median = 4 vs. 3; p <0.001). A total of 72% of weekends also had decreased numbers of CNAs assigned to the unit. Overall the number of CNAs was greater on weekdays than at weekends (median = 6 vs. 4; p <0.001). There were a total of 5 falls, one of which occurred on a weekday and 4 of which occurred at weekends. Only 1 of the falls at weekends had an associated injury. The rate of falls was therefore 1 (0.19%) of 522 weekdays as compared with 4 (0.77%) of 522 weekends (p <0.05). Only 2 omitted medications or medication errors occurred, both at weekends.
Quality appraisal 1 Case mix adjustment 2 Other adjustment 3 Uniform data collection 4 Participant follow-up 5 Random sampling 6 Geographical dispersal	 1 None stated 2 None stated 3 Unsure 4 No 5 No 6 One unit in Atlanta
Commentary	This facility typically staffed the subacute unit with full-time LPNs and CNAs every other weekend and on 3 weekdays. LPNs had received no special training. Staffing levels were not verified at the end of the shift; therefore the authors could not ascertain whether any staff were 'floated' to other floors. To avoid introducing bias into the outcomes, the staff were not informed of the results of the data recording. Although fewer LPNs and CNAs were on duty for the vast majority of weekends, staffing was not reduced at all weekends. The sample size is relatively small. The authors did not calculate a fall risk index for residents on the basis of the number of residents with chronic disabilities as others have done.
Research implications	Are current staffing levels adequate or detrimental to providing high-quality care to long-term residents?

This document was published by the National Coordinating Centre for the Service Delivery and Organisation (NCCSDO) research programme, managed by the London School of Hygiene & Tropical Medicine.

The management of the Service Delivery and Organisation (SDO) programme has now transferred to the National Institute for Health Research Evaluations, Trials and Studies Coordinating Centre (NETSCC) based at the University of Southampton. Prior to April 2009, NETSCC had no involvement in the commissioning or production of this document and therefore we may not be able to comment on the background or technical detail of this document. Should you have any queries please contact sdo@southampton.ac.uk.