SDO Protocol - project ref: 09/1006/25
Version: 1
Date: 10/08/2011

Using clinical practice variations as a method for commissioners and clinicians to identify and prioritise opportunities for disinvestment in health care.

Chief investigator	Dr. William Hollingworth
Sponsor	University of Bristol
Funder	SDO Programme
NIHR Portfolio number	Pending
ISRCTN registration (if applicable)	N/A

Table of contents

Using clinical practice variations as a method for commissioners and clinicians to identify and prioritise opportunities for disinvestment in health care.

Project Management

Project management team

The project team will consist of

Dr William Hollingworth, Reader in Health Economics, University of Bristol
Louise Tranmer, Co-Director of Commissioning, Bristol PCT
Dr. Chris Hine, Consultant in Public Health ,South Gloucester & Bristol PCTs
Dr. Padmanabhan Badrinath, Consultant in Public Health, Suffolk PCT
Prof. Jenny Donovan, Professor of Social Medicine, University of Bristol
Dr. Amanda Owen-Smith, Research Associate, University of Bristol
Prof. Jonathan Sterne, Professor of Medical Statistics and Epidemiology, University of Bristol
Dr. Penny Whiting, Senior Research Fellow, University of Bristol
Dr. Hayley Jones, Research Fellow, University of Bristol
Leila Rooshenas, Research Associate, University of Bristol
Theresa Moore, Research Associate, University of Bristol
John Busby, Research Assistant, University of Bristol

Day to day management

WH will line manage JB and will be responsible for the overall project budget and communication with the NIHR SDO. JS will supervise both JB and HJ in the analysis of HES data. PW will lead the technology appraisal and line manage TM. JD and AOS will lead the qualitative analysis and line manage LR.

CH and PB will represent the respective PCTs on the project team, co-ordinate collaboration with the PCT health analyst and knowledge officer and ensure access to PCT CPG/CAF committees.

The project management team will meet quarterly, in person or via teleconference to review study progress.

2. Aims/Objectives

The overall aim of this project is to develop and refine the process by which NHS commissioners identify clinical procedures that are over-utilised in their area and establish a process for disinvestment.

Our specific objective are:

1) To use routine data (HES) to evaluate the effectiveness of existing PCT commissioning criteria (issued during 2006-2008) in reducing the volume of procedures of uncertain clinical value.

2) To use routine inpatient data (Hospital Episode Statistics - HES) to identify procedures with the highest inter-PCT variation in use. High inter-PCT variance will be used as a proxy measure to identify procedures where there is likely to be uncertainty about the clinical value of the procedure in some patient subgroups (i.e. 'procedures of uncertain clinical value').

3) To work with two PCT commissioning groups to select two procedures of uncertain clinical value that they consider might be over-utilised by their local NHS trusts.

4) To conduct rapid technology appraisals and assemble national and local guidelines for these procedures to summarise the current evidence on effectiveness and cost-effectiveness. This evidence will be presented to and discussed with commissioning groups and local stakeholders who will, where appropriate, develop a disinvestment plan.

5) To use qualitative research methods to understand obstacles and solutions to the process of decision- making developed in each PCT for the proposed disinvestment through:

a) Observation of meetings between PCT commissioners and local primary and secondary care doctors and other stakeholders such as patient groups as they proceed with disinvestment.

b) Interviews with individuals involved in the decision-making process.

3. Background:

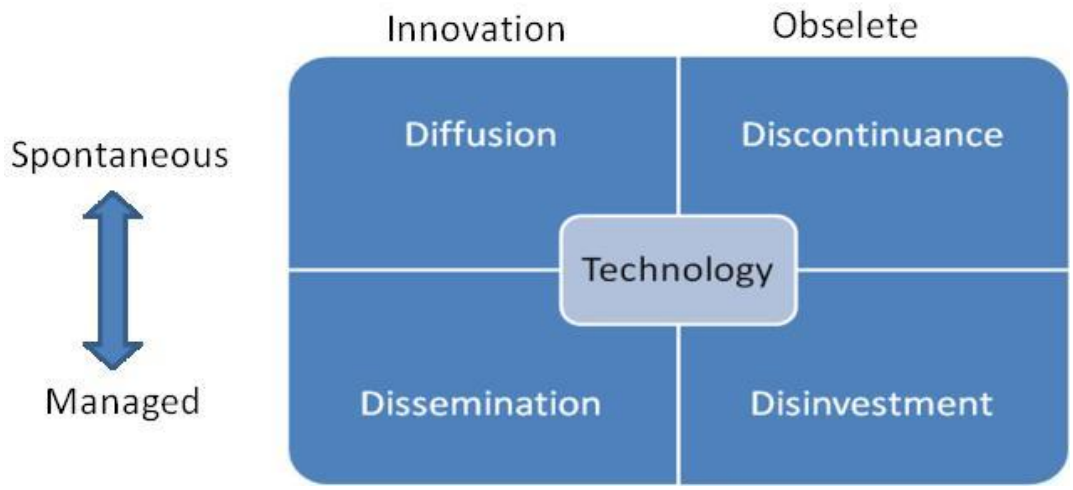
This section begins with a brief overview of the theory underlying technology diffusion and discontinuance and a discussion of the barriers to optimal disinvestment from obsolete medical technologies. In this project we use inter-PCT variation in procedure rates to identify medical technologies where there is professional uncertainty about the diagnostic threshold for a procedure or the value of the procedure in particular patient groups. Therefore, we next provide a description of the 'professional uncertainty hypothesis'. We then provide an overview of current nationwide and PCT initiatives to achieve disinvestment.

Theories of technology diffusion and discontinuance.

Rogers identified seams of diffusion and discontinuance theory in anthropology, sociology, economics, communication, and marketing.[1] Discontinuance of inefficient or inappropriately applied technologies will depend on characteristics of the technology (e.g. perceived relative disadvantage), characteristics of individuals

who use it (e.g. training and receptiveness to change), systems within which they operate (e.g. financial incentives) and interactions among each component. Rogers distinguishes between replacement discontinuance, which occurs when more efficient technology displaces the existing technology (e.g. CT replacing skull radiography in head trauma) and disenchantment discontinuance, which results when new information indicates that the benefits of the existing technology do not justify the costs or adverse effects (e.g. withdrawal of rofecoxib due to evidence on an increased risk of cardiovascular events).

Discontinuance can be spontaneous or managed (i.e. disinvestment – see figure). Reliance on spontaneous discontinuance will fail if there are imperfections in the 'market' for health care. In particular, imperfect evidence about the costs, effects and safety of existing interventions or lack of communication of this evidence to clinicians and patients will delay optimal discontinuance. Antman et al[2] reported that the majority of clinical experts recommended lidocaine for prophylaxis against ventricular fibrillation over a 25-year period despite successive trials showing no evidence of mortality reduction. Given this and other high profile examples of slow discontinuance, it is argued that NICE and PCTs should take a more proactive role in managing disinvestment to complement their existing work on managing the dissemination of new health technologies.



Overview of discontinuance and disinvestment in the health sector

Very little is known about the rate of health technology discontinuance or factors that facilitate it. In a review of more than 200 studies, Greenhalgh et al identified only one study that explicitly and prospectively studied discontinuance.[3] Therefore whilst there is growing recognition of the importance of disinvestment, there is little theoretical or empirical evidence to inform the disinvestment process. Pioneering work by Elshaug and colleagues[4-6] has identified five key challenges to health care disinvestment. 1) lack of resources to support disinvestment policy mechanisms; 2) lack of methods to identify and prioritise technologies with uncertain cost-effectiveness; 3) political, clinical and social challenges to changing established

practice; 4) lack of evidence on the efficiency of many existing technologies; 5) lack of funding for research into disinvestment methods.

This project is based on the hypothesis that high geographic variation in clinical procedure rates is an indicator of interventions where clinicians are uncertain of the diagnostic threshold or the clinical value in particular patient groups and therefore may be using the procedure inappropriately or inefficiently. As PCTs can easily benchmark procedure rates this is a potentially valuable method of addressing the second key challenge described by Elshaug. The qualitative component of our project will explore some of the barriers to disinvestment emphasised by Elshaug's third key challenge.

Geographic variation in procedure rates and clinical uncertainty

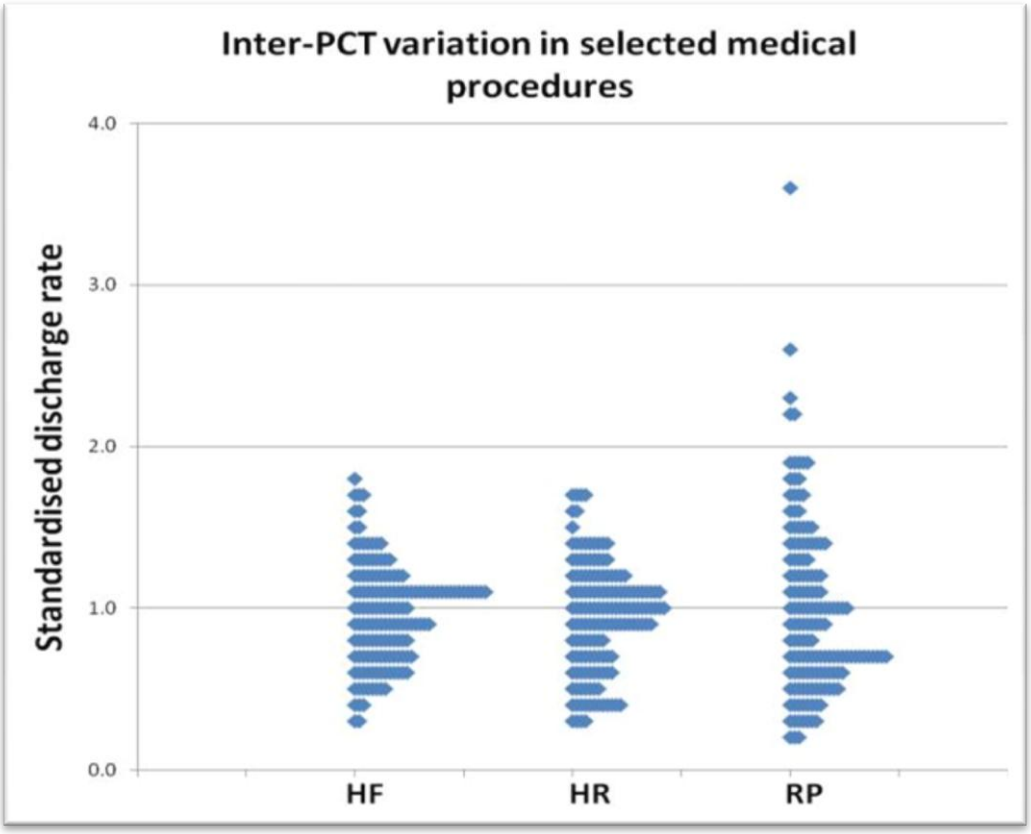
Glover, comparing pre-1945 tonsillectomy rates found such high geographic variations that he concluded that it was 'a prophylactic ritual carried out for no particular reason with no particular result'.[7] Wennberg has developed this into the 'professional uncertainty hypothesis'.[8] That is the theory that geographic variations occur because of differences among physicians in their diagnostic thresholds or in their belief in the value of the procedure, rather than any differences in clinical need. For example, hip fracture repair, where the diagnosis is clear cut and consensus on the value of surgery is high, has a low geographic coefficient of variation (CV=0.12) between regions of the US. In contrast, for lumbar spine fusion surgery where there is less agreement on the indications for surgery or the benefit of surgery, the CV is much higher (0.50).[9]

Geographic variation remains after adjustment for demographic factors and is unlikely to be due to differences in disease prevalence or patient preferences. Small area variations are also prevalent in the UK.[10] It is thought that variations build up over time as clinicians arriving in a region conform to existing practice patterns, due to local opinion leaders and educational fora that generate enthusiasm (or lack thereof) for a procedure.[11] Bisset[12] observed that as Scottish appendicectomy rates declined (2.89/1,000 in 1973 to 1.47/1,000 in 1993), there was a concurrent decrease in the amount of variation in procedure rates (0.25 coefficient of variation in 1973 to 0.16 in 1993) between the 12 Health Boards. She concluded that the reduced variation '… supports the view that improved management policies may have helped reduce 'professional uncertainty', unnecessary operations and variation in surgical practice.' Where there is marked practice variation, there is potential for evidence synthesis to inform commissioning criteria to standardise practice.

Pilot work exploring the use of HES data to identify potentially over-used procedures in the NHS

Most current data on clinical practice variation is derived from the Dartmouth Atlas project that compares rates of common surgical procedures among US Medicare patients from 306 referral regions. This suggests that procedures can be categorised

as ‘low variation’, ‘high variation’ and ‘very high variation’ based on the degree of heterogeneity in procedure rates.[13] In order to preliminarily evaluate the extent to which geographic variation in procedure rates is currently replicated among English PCTs, we selected three clinical procedures that, in the US, are categorised as ‘low variation’ (hip fracture repair (HF)), ‘high variation’ (hip replacement (HR)), and ‘very high variation’ (radical prostatectomy (RP)). The data (figure) represent crude procedure rates per capita among 152 English PCTs in 2007/8 based on hospital episode statistics (HES) returns. These pilot data confirm that geographic variation in English PCTs is also highest for radical prostatectomy (coefficient of variation (CV) of hip repair = 0.31, hip replacement = 0.33, radical prostatectomy = 0.53), although there is little difference in geographic variation between hip fracture and hip replacement. Five PCTs had radical prostatectomy rates more than twice the national average; one PCT exceeded three times the national average. Two of these are contiguous PCTs in the West Midlands and two are contiguous PCTs in South London suggesting that high procedure use is not contained within PCT boundaries.



Potential innocuous explanations for variation include: 1) expected statistical variation; 2) differences in clinical need 3) the effect of private health care provision; and 4) coding inaccuracies. We will identify procedures with rates that are more variable than can be explained by chance alone and rank procedure variability based on statistical criteria. We will also adjust for proxy measures of clinical need (e.g. PCT population age, sex, deprivation, and chronic disease prevalence) and the availability of private medical care (e.g. regional measures of access to private medical

insurance). A more detailed description of the statistical approach is provided in the methods section. Coding accuracy can be verified by PCTs and their local acute trusts. Other potential sources of variation in secondary care procedure rates are more informative for PCTs: 1) A shift from secondary to primary care (e.g. primary care vasectomy services); 2) Shift towards substitute procedures (e.g. hysterectomy via abdominal versus vaginal route) which may be more cost-effective; 3) Variation in GP testing and referral practices (e.g. use of prostate specific antigen (PSA) testing in asymptomatic men); and 4) Variation in specialists' threshold for treatment.

Should the NHS strive to eradicate all unexplained variation?

If there is compelling evidence that the procedure is no longer cost-effective in some or all of the clinical subgroups in which it is used, then there is a clear case for PCTs with high procedure rates to implement service changes to disinvest and standardise procedure rates. However, it has been argued[14] that when evidence is not compelling variation may represent legitimate differences in patient preferences for a procedure or merely unimportant eclectic practice by clinicians. Patient preference for surgery rather than more conservative care might explain some additional prostatectomy procedures in the five PCTs with more than double the national average surgery rate. However, it is unlikely to be the only explanation as research suggests that Urologists focus more on clinical parameters rather than patient views when recommending prostatectomy.[15] In some cases, variation due to eclectic clinical practice may be preferable to an enforced consensus around the prevailing, non-evidence-based, norm.[16] This variation might subsequently be used for observational comparisons of patient costs and outcomes. However, inpatient procedures tend to be both costly and invasive. Therefore, it seems reasonable that the onus should be on clinicians with high procedure rates to demonstrate (by actively participating in RCTs) that their approach results in better patient outcomes at an acceptable cost, rather than on the PCT to demonstrate otherwise.

A corollary of the discussion above is that PCTs with very low rates of procedures where evidence on cost-effectiveness is well established, might wish to consider whether they are providing optimum care for their population or whether they should invest more. This issue, while potentially interesting, is not the focus of this application.

Current disinvestment initiatives at the national level

The National Institute for Health and Clinical Excellence (NICE) was established to provide 'guidance on the use of new and *existing* medicines, treatments and procedures'. By 2006, NICE had published 113 technology appraisals. However, only two (wisdom teeth extraction & electroconvulsive therapy) evaluated existing technology rather than innovations. Furthermore, Sheldon et al[17] found that NICE guidance on wisdom teeth extraction in 2000 did not accelerate the already sharp decline in procedure rates that had been observed since 1996. Unlike emerging health technologies, there is rarely an industry or patient group pressing for

disinvestment. This preferential appraisal of new technologies has led to accusations that NICE is a 'golden goose' approving expensive new technology without fully considering whether savings opportunities can be realized through disinvestment in inefficient technologies.[18]

In 2006 NICE launched a new programme to 'reduce spending on [existing] treatments that do not improve patient care'. This programme faced immediate difficulties as 'in conversations with its stakeholders, NICE has received enthusiastic backing for the idea of appraising existing technologies to seek opportunities for disinvestment; but, when followed by requests for specific suggestions, the subsequent silence has been striking.'[19] In 2007, NICE was strongly criticised by the House of Commons Health Committee who recommended '… that more be done to encourage disinvestment. No evaluation of older, possibly cost ineffective therapies has taken place to date; … Our predecessor committee made the same recommendation; it is not acceptable that NICE continues to ignore this recommendation.'

Since then, the NICE disinvestment programme has evolved to include: 1) 'Recommendation reminders' to reduce ineffective care (e.g. 'Coronary revascularisation should not be routinely considered in patients with heart failure due to systolic left ventricular impairment, unless they have refractory angina.') drawn from existing NICE clinical guidelines; 2) 'Cost saving guidance' (e.g. computer rather than therapist delivered CBT) to identify more efficient ways of providing care; 3) 'Spending to save' to increase investment in public health interventions and save money through disease prevention; and 4) Commissioning tools and programme budgeting to help PCTs benchmark their service levels against clinical guidelines and each other.

It is too early to determine how effective these initiatives have been at achieving disinvestment. National disinvestment recommendations will only be successful if local clinicians and commissioners are aware of them, agree with them and act on them. PCT feedback at the 2009 Quality and Productivity Workshop organised by NICE, suggested that more needs to be done in order to realise this aspiration. Furthermore, other, local, initiatives are needed to complement the NICE pathways to disinvestment.

Current disinvestment initiatives in PCTs

Three PCTs have road-tested programme budgeting and marginal analysis (PBMA) as a tool for reallocating resources.[20] PBMA involves multi-disciplinary teams identifying options for change (both investment 'wish list' and disinvestment 'hit list'), scoring and prioritizing options and deciding on changes (often under the constraint of budget neutrality). Norfolk PCT identified four options for cutting back services, while Hull PCT (1 option) and Newcastle PCT (0 options) were less successful. These disinvestment options often dealt with local service configuration (e.g. reducing inpatient beds) rather than health technologies and it remains unclear whether this

pilot PBMA exercise resulted in any actual disinvestments.

An alternative, adopted by many PCTs, is to set up criteria based access (CBA) or low priority procedure lists that limit access to procedures of questionable clinical and cost-effectiveness. For instance, in the 'Save to Invest' programme, South West London PCTs generated a list of 34 'relatively ineffective'; 'largely cosmetic'; 'questionably effective in mild cases'; or 'effective only after failure of 1st line therapy' procedures. Standardised referral criteria were agreed and it is estimated that pan-London implementation could save £93 million. A key barrier to the wider use of this approach is the lack of a standard methodology for identifying locally over-utilised procedures. Therefore, the selection of procedures by PCTs is fairly arbitrary and predictable (e.g. aesthetic procedures).

Existing free inpatient benchmarking tools (e.g. NHS comparators, PbR benchmarker) are based on Healthcare Resource Groups (HRGs) which 'bundle' together many different procedures and are therefore not well-suited to identifying individual over-utilised procedures. Proprietary benchmarking tools (e.g. Dr Foster) allow PCTs to drill down to individual procedures, but not all PCTs subscribe. In this proposal, we aim to develop a sustainable method for PCTs to use routine data on variation in procedure rates as a tool for identifying potential health technologies

4. Need

UK health care expenditures have risen from £54bn (6.6% of GDP 1997) to £109bn (8.4% of GDP 2006), but increases will not continue indefinitely. A recent King's Fund report concluded that there was likely to be little real growth in NHS expenditure between 2011/12 and 2016/17. The March 18th 2010 issue of the BMJ contained an editorial and four linked articles on disinvestment in health care.[21-24] They identify a plethora of ideas for achieving disinvestment including better alignment of primary and secondary care, better integration of the health, social and community care systems, and avoidance of ineffective or inefficient medical tests and procedures. But this is tempered by the recognition that we 'tend to underestimate the factors that promote resistance to the kinds of change a strategy of disinvestment is bound to cause.'

Technology is thought to be a primary cause of increasing healthcare costs. The NHS has invested heavily in methods and processes (e.g. NICE) to ensure that new health innovations are safe and cost-effective before diffusion. However, methods for identifying and disinvesting from existing technology that has been discovered to be inefficient, ineffective or harmful are less developed. Therefore Primary Care Trusts (PCTs), who have a mandate to fund NICE-approved technology appraisals, often struggle to balance budgets. PCTs have a key role in identifying interventions which are currently over-used in their area, assembling the evidence about appropriate cost-effective use and negotiating with clinical and other local stakeholders to rationalise service provision. World Class Commissioning (WCC) challenges PCTs to work with the public, GPs (practice based commissioners) and secondary care

clinicians to prioritise investment such that 'investment decisions are made in an informed and considered way, ensuring that improvements are delivered within available resources'. Previous work on selecting procedures for potential disinvestment has often relied on clinician introspection to identify topics for review. This has generally resulted in a list of 'usual suspects' (e.g. tonsillectomy, grommets for otitis media), but these do not always have local relevance and this method is unlikely to be sustainable in continuously identifying locally over-utilised procedures in successive commissioning cycles.

The aim of this proposal is to develop a method for PCTs (and, in the future, clinical commissioning groups) that will involve clinicians, patients, and academic partners and use geographic variance in procedure rates, benchmarking, evidence synthesis and consensus building to identify and prioritise technologies for disinvestment. Disinvestment does not necessarily mean the cessation of services, but could take the form of contraction through new or adapted referral guidelines for GPs, reduced indications for procedures agreed by clinicians, or a shift from higher to lower cost service settings. The McKinsey Report estimated that such allocative efficiencies could yield up to £6.6bn in savings across the NHS.[25] Unless PCTs find successful ways to disinvest from inefficient health technologies, it will be impossible for the NHS to continue to afford drug, device and service innovations that improve the health of the population.

5. Methods

Setting

The research team will work with two PCT commissioning groups. NHS Suffolk (Suffolk PCT) is a large PCT with a registered population of over half a million and is predominantly rural. The PCT commissions services from two district general hospitals and is also served by a Foundation Trust. The PCT has well established Clinical Prioritisation system and process which enjoys broad support across the health economy. The Clinical Prioritisation function in the PCT is led by the Public Health Directorate and Suffolk PCT's Public Health Department is one of the strongest in the country with nine consultants. The PCT also has an interest in health services research with two of its consultants holding doctoral degrees and academic appointments.

NHS Bristol, NHS North Somerset and NHS South Gloucestershire (BNSSG) refers to a health community of 3 PCTs with a total population of 885,000. The PCTs have a joint Commissioning Advisory Forum to develop commissioning policies on single health care interventions. These are identified from requests made to exceptional funding panels, direct applications by local NHS trusts and PCT staff. The interventions can be new or established, but have questions arising about appropriate commissioning policy. Following consideration of evidence and policy options, CAF recommends a policy for adoption to the Professional Executive Committees (PEC) of the 3 PCTs. The CAF is led by NHS Bristol. Policy adoption and implementation rests with the local

PCT. CAF membership includes clinical and management representatives from local NHS providers, PCT commissioning management, public health and GP PEC chairs.

Our research partnership provides a mix of a regional urban centre PCT with a more ethnically diverse population and pockets of high deprivation (Bristol) and PCTs serving more rural, wealthier and predominantly white populations (Suffolk, N Somerset, S Gloucs). Both PCT commissioning groups already use criteria based access and case-by-case individual funding panel approval, based on exceptionality, to guide use of low priority procedures. We have established many links between the research team and the PCTs, including full support from the PCT CEOs (letters of support available on request), PCT directors of commissioning and consultants in public health as co-applicants and funded time for PCT health analyst and knowledge officers to work with the academic teams on this research project. These multi-layered links will ensure that the research project is a true collaboration.

The 2010 Government White Paper, 'Liberating the NHS: Commissioning for Patients' sets out proposals for putting local consortia of GPs and other clinicians (now termed clinical commissioning groups (CCG)) in charge of commissioning services to best meet the needs of local people, supported by an independent NHS Commissioning Board and local authority Health and Wellbeing Boards. The timetable outlined in the White Paper indicates that CCG will be introduced in shadow form during the first 15 months of this research project, becoming formally established with indicative budgets in 2012/13. Geographical variation and local high-utilisation of secondary care procedures will continue to be a key indicator of potential areas for disinvestment as commissioners transition from PCTs to CCGs. Three of the 'Pathfinder' groups, named in December 2010 lie within our partner PCTs (South Glos Consortium Ltd, East Suffolk Federation, Ipscom, Ipswich). Therefore, we will be well placed to tailor our research to meet the evolving needs of CCGs and also include key CCG staff in the qualitative research.

Objective 1: *To use routine data (HES) to evaluate the effectiveness of existing PCT commissioning criteria (issued during 2006-2008) in reducing the volume of procedures of uncertain clinical value.*

Design

We will investigate the impact of PCT commissioning criteria for five procedures based on the following criteria: 1) there is variance between the two PCT commissioning groups (i.e. policy versus no policy); 2) the PCT policy was implemented before Jan 2009; 3) the procedure has high volumes; 4) there is no NICE technology appraisal on the procedure. We will work with both PCT commissioning groups to verify the date when the policy was implemented and ensure that there was no previous policy. We will contact other PCT public health teams in the South West and East of England regions to ascertain their commissioning policies on the selected procedures with a view to including them in the analyses.

Data collection

Many PCTs have published criteria for restricting the use of 'low priority' procedures. The table compares the published policies of Suffolk and Bristol, North Somerset and South Gloucestershire [BNSSG] PCTs. There are some similarities in the procedures selected by PCTs (particularly aesthetic procedures such as abdominoplasty and face lifts), however there is inter-PCT variance in the date at which threshold policies were introduced and, in many cases, the procedures that have criteria based access. For example Suffolk PCT have issued a threshold policy for lumbar spine fusion surgery, whereas BNSSG PCTs have a policy for knee arthroscopy. This variation provides a natural experiment to evaluate the effectiveness of PCT commissioning criteria in preventing inappropriate use of low priority procedures.

Procedure	**Suffolk PCT**	**BNSSG PCTs**
Abdominoplasty	Dec 2006	Nov 2008
Bariatric surgery		June 2008
Benign Skin Lesions	Dec 2006	Mar 2009
Bone anchored hearing aids and cochlear implants		Sep 2008
Brachytherapy, Cryotherapy, HIFU for prostate Ca.		Undated
Breast Augmentation	Dec 2006	Sep 2009
Breast Reduction	Dec 2006	June 2009
Buttock lift	Dec 2006	Jan 2010
Cataract Surgery	Dec 2006	Pending
Cryoablation of renal mass		Sep 2009
Deep brain stimulation		Mar 2007
Dental Titanium Implants	Dec 2006	May 2008
Dilatation & Curettage and Hysteroscopy	Jan 2007	
Excimer Eye Laser Surgery		Jan 2010
Face lifts	Dec 2006	Sept 2009
Ganglion, CTS, Dupuytrens, Trigger finger	Dec 2006	Sep 2008
Grommets for Otitis Media with Effusion	Dec 2006	
Hip and knee replacement surgery	June 2007	
Hysterectomy for Heavy Menstrual Bleeding	Dec 2006	
Infertility [IVF/ICSI/IUI]	Dec 2006	Jan 2010
Instrumented gait analysis		Mar 2003
Knee arthroscopy		Mar 2008
Labiaplasty, vaginoplasty and hymenorrhaphy		Mar 2009
Laser lumbar surgery	Dec 2006	
Lenses for Scotopic Sensitivity Syndrome	Dec 2006	
Liposuction	Dec 2006	Jan 2010
Lumbar spine fusion & discectomy	May 2007	
Male Circumcision	Dec 2006	
Mastopexy	Dec 2006	Mar 2009
Occipital nerve stimulation		Sep 2008
Orthodontics	Dec 2006	
PDT for Age-related MD	Dec 2006	

Percutaneous aortic/pulmonary valve replacement		June 2008
Pinnaplasty	Dec 2006	June 2008
Reduction of ear lobes	Dec 2006	
Reversal of female sterilisation	Dec 2006	Jan 2010

Data analysis

Denoting the observed and expected number of times a particular procedure was used in PCT i in year t by O_{it} and E_{it} respectively (where the E_{it}'s have been calculated to take into account proxy measures of clinical need, as described previously), we will use the following Poisson regression analyses to evaluate the effect of PCT commissioning guidelines, based on annual data from 1999/2000 – 2009/10. These analyses are known in econometrics as difference in differences analyses:

$$O_{it} \sim \text{Poisson}(r_{it}\, E_{it})$$
$$\log(r_{it}) = \beta_0 + \beta_1 C(t) + \beta_2\, i + \beta_3\, (C(t)\, .\, i)$$

where i = 0, 1; t = 1,..., 11; and C(t) = 1 if the commissioning criteria were in force during year t, and 0 otherwise. The coefficient β_3 measures the association of procedure rates with introduction of the commissioning criteria, after adjusting for the time trend in PCTs without any commissioning criteria for that procedure (β_1) and baseline differences in procedure utilisation (β_2) between PCTs.

In provisional analysis, we will use the same statistical approach to evaluate whether the disinvestment commissioning criteria that are developed as a result of this project led to an actual change in practice. In order to conduct this analysis, we will use provisional monthly HES data to monitor procedure volumes in the 12 months pre- (approximately Jan 2011 – Dec 2011) and 6 months post- (Jan 2012 – June 2012) disinvestment commissioning criteria. This will provide the PCTs with an initial estimate of whether the disinvestment process has resulted in savings due to a reduction in inpatient procedures.

Objective 2: *To use routine inpatient data (Hospital Episode Statistics - HES) to identify procedures with the highest inter-PCT variation in use. High inter-PCT variance will be used as a proxy measure to identify procedures where there is likely to be uncertainty about the clinical value of the procedure in some patient subgroups (i.e. 'procedures of uncertain clinical value').*

Design

We will use national HES data to identify high-volume, high variance inpatient procedures. HES is the national statistical data warehouse for England tracking inpatient and day case care provided in NHS hospitals and for NHS patients treated in independent hospitals. HES data are available from 1990 onwards and contain information on diagnoses, procedures and demographic and geographical data on the patients treated.

Procedures can be tracked in HES using OPCS codes. OPCS 4.4 defines 7,212 specific procedures (e.g. M61.3 'Transvesical prostatectomy') grouped into 1,278 major procedure groups (e.g. M61 'Open excision of prostate') and 23 chapters (e.g. M = 'Urinary'). PCTs might wish to benchmark their performance on the 23 procedure chapters to identify programmes of care where they are outliers. However our aim is to identify procedures for appraisal and potential disinvestment. Therefore, we will focus our analysis on the 1,278 major procedure groups. HES contains up to 24 procedure fields for each finished consultant episode. We will include all procedures performed during an inpatient episode, not only the procedure identified as the main procedure. An inpatient spell may consist of multiple episodes if, for example, the patients care is transferred to another consultant. We will use data from the first episode to exclude double counting of procedures.

In 2008/9, 271 (21%) of the major procedure group codes accounted for 8.4 million (90%) inpatient procedures. Over a one year period, a typical PCT would commission between 50 and 5,500 of these most frequently performed procedures. By focusing on these most frequently used procedures we will identify those that are most economically influential for the local PCT and where geographic variation in utilization can be quantified precisely. Preliminary analysis revealed implausible utilisation rates and variation between many miscellaneous and diagnostic procedures (e.g. tuberculosis support, Injection of radiocontrast material) which are thought likely to be caused by inconsistent coding across PCTs. This spurious variation will be avoided by excluding all diagnostic and miscellaneous procedures hence our analysis will focus on only interventional procedures. Therefore, our analysis includes 190 procedures.

We will use annual rather than quarterly HES data to exclude temporary spikes in procedure rates (e.g. to meet waiting list targets or due to temporary ward closures). Geographic categorisations will be based on the patient's PCT of residence rather than the PCT of treatment, to avoid distortion due to cross-PCT referrals for care. In order to gauge the stability of high-variance procedures over time and to allow PCTs to track their performance over time, we will conduct the analysis stratified by the three most recent years of HES data (2007/8, 2008/9, 2009/10). In the small minority of acute NHS trusts where HES data quality notes indicate coverage issues (i.e. missing episode records) for any month(s) during the year, we may conduct sensitivity analysis initially excluding data from these trusts and then imputing procedure rates for these trusts based on procedures performed during the rest of the year and previous years. If the identification of high variance procedures is not robust to these sensitivity analysis, we will use more complex methods (i.e. multiple imputation) to calculate utilisation rates. For each PCT we will calculate annual utilisation procedure rates. 'Expected' utilisation rates based on proxy measures of clinical need ('case mix') will also be calculated based on ONS mid-year estimates of PCT population. As described by Ibanez et al,[26] the ratio of observed to expected rates then provides a case-mix adjusted measure of utilisation in each PCT.

In all analyses we will include PCT-level estimates of population, age, gender,

deprivation, ethnicity, prevalence of chronic disease (asthma, atrial fibrillation, CHD, CKD, dementia, diabetes, hypertension, stroke, all cause cancer) and lifestyle factors (smoking, binge drinking, obesity) as covariates to adjust for the role that these factors play in determining clinical need. Other things being equal, we would expect PCTs in areas with widespread access to private medical services to have lower rates of procedures performed in NHS hospitals or on NHS patients in independent hospitals. Analysis of coronary revascularisation in London suggests that 12% of procedures were performed on privately insured patients in private hospitals and therefore would be excluded from HES data.[27] Data from the British Household Panel Survey (BHPS) indicate that over one quarter of residents in outer London are covered by private medical insurance in their own name (18%) or via another family member (10%). Whereas, in Tyne and Wear, 6% of residents had private medical insurance in their own name and 2% had coverage via a family member. In our analyses we will use these region-level estimates of private medical insurance coverage as a proxy for privately-funded procedures performed in the independent sector. PCT-level deprivation indices are negatively associated with the volume of privately-funded procedures,[27] therefore including PCT deprivation scores as a covariate in analyses will adjust for some sub-regional variations in access to private medical care.

Data Collection

The University of Bristol, Department of Social Medicine holds HES inpatient data from 1991/2 onwards, with ethical approval subject to Section 251 of the NHS Act 2006. We have submitted our annual review documentation to the Ethics & Confidentiality Committee of the National Information Governance Board for Health and Social Care. HES data for the financial year 2009/10 will be purchased from the NHS information centre in December 2010.

Data analysis

Naïve measures of variation in rates include the external quotient (the ratio of the lowest to the highest observed rate)[28,29] and the coefficient of variation (CV), which is based on a measure of observed variability. However, it is now well recognised that these measures are inappropriately influenced by random fluctuations.[30] More appropriate measures, which separate the 'true' or 'systematic' between-area variability from sampling error can be derived based on a hierarchical model structure.[26,31] We will therefore fit separate random effects Poisson models to the utilisation rates for each procedure, adjusting for case-mix by including the logarithm of the expected counts as offsets in the regressions. An appropriate estimate of true variability in the utilisation of each procedure will be obtained based on these models, along with measures of the associated uncertainty.

To give an indication of which procedures might have most variation in use, the point estimates of variability for each of the 190 major procedures can be ranked. However, such rankings ignore the uncertainty associated with each measurement and

therefore have a tendency to draw attention to unimportant differences.[32,33] We will therefore define procedures as 'very high' variance if there was greater than a 0.95 probability that inter-PCT variation in the use of that procedure was more than three times higher than the median inter-PCT variation for all 190 interventional procedures. Similarly we identified 'high' (2-3 times), 'average' (0.5-2 times), 'low' (0.33-0.5 times) and 'very low' (<0.33 times) variance procedures. These results will be available within 6 months of project commencement. As CCGs become established and geographic boundaries are defined it will be relatively straightforward to replicate our analyses to identify inter-CCG variation. Contingent upon CCG boundaries being established by June 2012, we will use patient postcodes to map procedure use by CCG and publish our findings during the last 6 months of the project.

If the variability measures are themselves very dispersed (a large number being significantly different from the average) and / or if some are very imprecise, then an additional hierarchical model will be fitted to the variance components themselves. This will lead to some *shrinkage* of each measure towards the average, which is well known to have appealing properties in a variety of contexts. [34,35] Hierarchical models for variance components have been suggested previously by Gelman.[36] Based on such a model, outlying or extreme procedures would then be identified using methods similar to those described by Jones and Spiegelhalter.[37]

Using the derived list of high variance procedures, we aim to further identify two of these procedures as being of high *local* use in our collaborating PCTs. Again, the methods discussed by Jones and Spiegelhalter[37] can be used here, to identify outlying or extreme health-care providers. We will develop guidelines which other PCTs can use to identify procedures which are potentially over-used in their local health economy.

Objectives 3&4: *To work with two PCT commissioning groups to select procedures of uncertain clinical value that they consider might be over-utilised by their local NHS trusts. To conduct rapid technology appraisals and assemble national and local guidelines for these procedures to summarise the current evidence on effectiveness and cost-effectiveness.*

Scoping the technology appraisal

Based on the results from objective 1, the research team will present a shortlist (selected from all 190 frequently used procedures) of potential technologies for review by the PCT clinical priorities group (Suffolk CPG) or commissioning advisory forum (BNSSG CAF). The exact criteria for selecting procedures will be developed with PCTs based on preliminary analysis. However, we expect that selected procedures will have high inter-PCT variance (suggesting that there is some discretionary use) and will have high local use. If necessary, secondary criteria such as procedures where

hospital length of stay for PCT patients is more than 1.5 times the median length of stay will be used to prioritise procedures for inclusion on the shortlist. The CPG/CAF will review the shortlist, conduct local validation to explore potential anomalies (e.g. review of medical records to identify coding errors by local trusts), search for any existing guidance (e.g. NICE) on the procedure and select two high variance high local use procedures for appraisal. We will then develop a scoping document for the technology appraisal (TA) detailing the patient group(s), the intervention, the comparison(s) and the outcome(s) of interest. The scope of the TA will be narrow to ensure that it can be achieved in a timely fashion to aid the commissioning decision.

Conducting the technology appraisal

A rapid systematic review of the literature will be undertaken in broad accordance with the Centre for Reviews and Dissemination (CRD) guidelines for undertaking systematic reviews and the Cochrane handbook for systematic reviews of interventions.

A database of published and unpublished literature will be assembled from systematic searches of electronic sources, hand searching, and consultation with experts in the field. Existing systematic reviews will be identified through searches of DARE and the Cochrane Library. If a recent, relevant, high quality review is identified, our technology appraisal will go no further than updating and interpreting the existing review for the PCT commissioners. If no such review is found, we will conduct a rapid systematic review. Rapid systematic reviews streamline the traditional systematic review processes in order to synthesize evidence in a shortened timeframe most useful to policy makers.[38] The process is streamlined by selecting a focussed research question, limiting the search to the most common electronic sources and restricting the review to the highest quality evidence (e.g. RCTs).

Primary studies will be identified by searching MEDLINE and EMBASE. Attempts to identify further studies will be made by examining the reference lists of all retrieved articles and contacting experts in the field. We will search for existing guidelines among NICE technology appraisals, clinical guidelines, public health guidance and the NHS National Library of Guidelines. We will contact the Regional Priorities Advisory Committee for any existing PCT commissioning criteria that have been issued on the procedure.

The results of the searches will be screened for relevance independently by two reviewers. Disagreements will be resolved through consensus or referral to a third reviewer where necessary. Studies that appear potentially relevant will be ordered and assessed for inclusion by one reviewer and checked by a second. Data extraction forms will be developed using Microsoft Access. These will be piloted on a small selection of studies and adjusted as necessary. Study data will be extracted by one reviewer and checked by a second. Disagreements will be resolved through consensus or referral to a third reviewer where necessary. Data will be extracted on the following: study details (identifier, study design, location, year), participant details

(number of participants, age, gender, other relevant details), intervention details, comparator details and effect sizes. For RCTs, we will use the Cochrane Collaboration risk of bias tool to assess the internal validity of the study results. Systematic reviews will be assessed for methodological quality using the DARE database inclusion criteria. The assessment will be carried out by one reviewer and checked by a second. Disagreements will be resolved through consensus or referral to a third reviewer where necessary.

The search strategy for identifying efficacy studies of the selected procedures will be modified to identify studies investigating the cost and outcomes of these procedures. The search strategy will be adapted to focus on economic evaluations using MEDLINE and Embase filters. In addition searches of the NHS Economic Evaluation Database (NHS EED) and Health Economic Evaluation Database (HEED) will also be carried out. The quality of primary economic evaluations will be assessed using the Quality of Health Economic Studies (QHES) instrument.

The draft results of the rapid technology appraisal will be presented to the CAF/CPG in December 2011 for development and feedback. At these meetings, the potential to disinvest from the selected procedures will be considered alongside all other proposals for PCT investment/disinvestment plans (i.e. investment business cases, uncontrolled activity growth estimates, NICE guidance). If, based on the technology appraisal, the selected procedures are considered to be over-utilised, the PCT public health and commissioning team will develop a 'threshold policy' (procedures for which there is a set of criteria that need to be met before the PCT will provide funding), or a 'partially excluded procedure policy' (procedures that will not normally be funded and require case-by-case individual funding panel approval based on exceptionality). By design, the procedures selected for appraisal will be frequently used, therefore it is more likely that the PCT will develop a 'threshold policy' to guide use rather than an exclusion policy. These threshold policies include a start date, planned review date, brief background information, eligibility criteria, rationale for the decision and references. For example, Suffolk PCTs current eligibility criteria for tonsillectomy are (sore throats due to tonsillitis AND five or more episodes of sore throat per year AND symptoms for at least a year AND the episodes of sore throat are disabling and prevent normal functioning).

Once the draft has been developed, this will be sent out for wider consultation to seek the views of all key stakeholders, modelled on the current NHS Suffolk process which has been used successfully for four years. A consultation pack will be produced which will include the draft policy, the rationale behind the policy including the evidence brief and a list of specific questions. The stakeholders will be asked whether they agree with the policy and any changes they would like to suggest and the rationale behind their suggestions. They will also be given an opportunity to add any further evidence not picked up by the policy development team. The stakeholders are lead clinicians in secondary care through Medical Directors, service level managers, GPs, patient groups, Practice Based Commissioners, PCT commissioning & finance colleagues, and the local medical committee. All the responses will be summarised

and presented along with the draft policy for deliberation before finalising the policy.

Objective 5: *To use qualitative research methods to understand obstacles and solutions to the process of decision-making developed in each PCT for the proposed disinvestment through (a) observation (and tape-recording where possible) of meetings between PCT commissioners and local primary and secondary care doctors and other stakeholders such as patient groups as they proceed with disinvestment, and (b) interviews with individuals involved in the decision-making process.*

Design

Qualitative research methods (observation and interviewing) will be used to understand the development and implementation of the disinvestment decision-making from the perspectives of commissioners, clinicians, patients and any other stakeholders. The ultimate aim of this part of the project will be to assist in the refinement of the process for future use through the identification of factors that facilitated service changes, and thus disinvestment, as well as identifying issues and approaches that acted as obstacles to development or implementation.

Data collection

Qualitative data collection will occur in both PCT prioritisation groups at several stages during the disinvestment process. First, in-depth semi-structured interviews will be undertaken by the researcher with PCT Chief Executive Officers and members of the Commissioning Advisory Forum (CAF – BNSSG) or the Clinical Priorities Group (CPG – NHS Suffolk). Eight to ten individuals will be purposively selected from each CAF/CPG to include informants across a range of roles. For example at NHS Suffolk this will include the director of public health, director of commissioning, a non-executive director, a consultant in public health, GP practice based commissioners (including the nascent CCG), a lay advisory group member, and the medical directors from the two local NHS hospital trusts. These interviews will occur before the list of high-variance high-use procedures is finalised. The aim of the interviews will be to elicit members' views on previous disinvestment initiatives, the effectiveness of existing commissioning criteria at controlling procedure rates, expectations for the disinvestment initiative proposed in this research project and the obstacles they think that the disinvestment process is likely to encounter. These data will provide essential background information related to the commitment of CAF/CPG members to the process and help identify factors that will need to be taken into account in planning the engagement with other stakeholders.

The researcher will observe key meetings of the CPG/CAF throughout the 2011/12 commissioning cycle. The exact timing of observations will be dependent on the scheduling of meetings and the agenda of those meetings. In NHS Suffolk, the CPG meets monthly, therefore we plan a preliminary visit in April 2011, to describe the study to the CPG members in more detail and request consent for participation in the qualitative study. Other key stages will include the presentation of results from the

HES analysis (objective 1) to the CPG in June 2011, presentation of the technology appraisals to the CPG in approximately November 2011, and discussion of draft commissioning criteria in approximately March 2012. The researcher will observe (and audio-record if possible) these key meetings and all interim CPG meetings where disinvestment proposals are on the agenda. CPG/CAF recommendations for investment and disinvestment are referred to the professional executive committee (PEC – BNSSG) or the clinical executive committee (CEC – Suffolk) for consideration. These PEC/CEC meetings discussing the two procedures selected for potential disinvestment will also be observed. Where it is not possible to record meetings, detailed field-notes will be taken.

After the draft commissioning policy has been sent to stakeholders for feedback, the third stage of the qualitative enquiry will involve in-depth interviews with representatives from as many of the stakeholder groups as possible. These will include service managers, primary care physicians, hospital consultants, and patients. Patient representatives will be identified by drawing on the lay groups established by the PCTs or those related to the hospital specialists. If no specialty patient groups exist, hospital consultants will be asked if they could identify suitable patients for interview. These interviews will explore the acceptability of the draft policy to managers, clinicians and patients, and will address a number of issues, including the potential conflict between restricting procedures and patient choice, and the types of information needed by patients to ensure decision-making processes are as acceptable as possible. All interviews will be audio-recorded, subject to participant consent.

At the end of the project, further interviews will be undertaken with PCT CEOs, members of the CPG/CAF and members of the new CCG. The aim of these interviews will be to allow informants to reflect on the disinvestment process and identify aspects that could be improved in future. These data will be used to make any adaptations required to the disinvestment process. The finalised disinvestment process will be published on the research project website so that it can be replicated by other PCTs/CCG.

Data analysis

The main analytic approach used will be coding and constant comparison, which will involve detailed scrutiny of the transcripts of interviews and meetings (and field notes taken during meetings it was not possible to record), which will then be coded to identify and inter-relate emerging themes.[39]

The initial aim of analysis will be to document how the disinvestment process worked in practice at the PCTs, and to identify those factors that enabled and disabled attempts at disinvestment. Additionally, an important analytic aim will be to assess the input of various stakeholder groups into the process of decision-making, and to consider to what extent this affected the eventual acceptability of draft guidance to patients and clinicians. This will be achieved through analysing observation and

interview data to understand the views of different stakeholder groups on what factors they felt were most important at different stages of the decision-making process. This information will then be compared to data arising in later interviews with clinicians and patients who were not involved in the disinvestment process. A further aspect of analytic interest will be the interaction between different stakeholder groups during the decision-making process. Informants' views on their input into the process will be explored at all stages of the process, and conversation analysis techniques will be used to examine the interactions between various stakeholders during CAF/CPG/PEC meetings.[40]

Detailed descriptive accounts of the data will be written, and matrices will be drawn up to facilitate the inter-relation of emergent themes, to make comparisons between the operation of the disinvestment process at the two sites, and to compare the views of different stakeholder groups.[41] Data analysis will run in parallel with data collection so that emerging themes of interest can be followed up, and particular attention will be paid to the emergence of 'negative cases' in the data, where an individual's view or experience is particularly divergent to the dominant view.

6. Contribution of existing research:

The research will contribute to the collective research effort by developing a blueprint that PCTs/CCGs can use for identification of procedures of uncertain clinical value.

Outputs

1) A ranking of high-volume, high-variance inpatient procedures. This ranking will be generated during the first phase of the project and will be available in *June 2011* and therefore able to inform 2011/12 PCT commissioning. The most valuable route of access for PCTs would be an online resource similar to the Dartmouth Atlas of Health Care (http://www.dartmouthatlas.org/index.shtm) which would allow individual them to benchmark their performance on all procedures. Full website development will not be possible within the resources in this application. However, we will develop a prototype online database and discuss further dissemination with existing NHS organisations (e.g. NHS comparators, PBA interactive atlas).

2) An evaluation of the impact of existing PCT commissioning guidelines/policies.

3) Four technology appraisals of procedures of uncertain clinical value. The primary target for these reports will be the two PCT commissioning groups, however other PCTs who also have high utilisation of these procedures would also benefit.

4) A qualitative exploration of PCT CAF/CPG and CCG members views on disinvestment and perceived barriers to it.

5) A qualitative description of the disinvestment process at the PCTs

6) A report and action plan to the NIHR SDO, including policy recommendations

We will disseminate our research during 2011-13 at conferences such as the HSRN / SDO Network annual conference and the NHS confederation annual conference. We will submit our work to policy journals (e.g. Health Service Research and Policy) and specialty journals related to the clinical areas identified for disinvestment to reach the wider academic and NHS communities. The research project will post the final proposed process for disinvestment on a website, so that commissioners from PCTs across the country can comment on, replicate and develop the process. We have discussed this project proposal with members of a PCT Lay Advisory Group to get feedback on the clarity of the proposal. If successful, we will work closely with lay collaborators from this group in finalising the protocol and interpreting and disseminating the results.

7. Plan of Investigation:

The investigation is conducted in three phases:

Phase 1 (Pre-start date –March 2013) : Quantitative analysis - HES data, impact of commissioning criteria
Appointment of research assistants
Prepare HES data extract
Analysis of variability
Identification of PCT high use procedures
Identification and analysis of effectiveness of existing criteria
Publication of quantitative analysis

Phase 2 (Sept 2011 – June 2013) : Conduct technology appraisals for selected procedures
Shortlist procedures and scope technology appraisals
Conduct technology appraisals
Publication of technology appraisals

Phase 3 (July 2011-March 2013) : Qualitative analysis of the 2011/2012 commissioning cycle
Request research consent
Initial interviews with GAF/CPG members
Observed CPG/CAF meetings
Interviews with stakeholders on draft criteria
Closing interviews with CAF/CPG members

8. Service users/public involvement

Service users have been involved in the development of this study and will be included on the advisory panel. The representatives on the project advisory panel will

be paid according to INVOLVE guidelines. In addition lay members of the CAF/CPG committees will be key informants for the qualitative study. Once specific procedures have been identified for technology appraisal, we will seek representatives from the most relevant patient groups via the NHS trust and PCT Patient Advice Liaison Service and by contacting local user groups. These representatives will then be invited to comment on draft commissioning criteria and participate in the qualitative study. One frequently cited barrier to disinvestment is the media and public pressure faced by PCTs when attempting to reduce the use of ineffective or inefficient care. An important aim of the qualitative research will therefore be to focus on the patient experience of the disinvestment process.

9. References:

Reference List

1. Rogers EM. Diffusion of Innovations. 5thed. New York, NY: The Free Press; 2003.

2. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 1992;**268**:240-8.

3. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. Milbank Q 2004;**82**:581-629.

4. Elshaug AG, Hiller JE, Tunis SR, Moss JR. Challenges in Australian policy processes for disinvestment from existing, ineffective health care practices. Aust.New Zealand.Health Policy 2007;**4**:23.

5. Elshaug AG, Hiller JE, Moss JR. Exploring policy-makers' perspectives on disinvestment from ineffective healthcare practices. Int.J.Technol.Assess.Health Care 2008;**24**:1-9.

6. Elshaug AG, Moss JR, Littlejohns P, Karnon J, Merlin TL, Hiller JE. Identifying existing health care services that do not provide value for money. Med.J.Aust. 2009;**190**:269-73.

7. Bevan G, Hollinghurst S, Benton P, Spark V, Sanderson H, Franklin D. Using information on variation in rates of supply to question professional discretion in public services. Financial Accountability and Management 2004;**20**:1-17.

8. Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problem of supplier-induced demand. Soc Sci Med 1982;**16**:811-24.

9. Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES. United States' trends and regional variations in lumbar spine surgery: 1992-2003. Spine 2006;**31**:2707-14.

10. McPherson K, Wennberg JE, Hovind OB, Clifford P. Small-area variations in the use of common surgical procedures: an international comparison of New England, England, and Norway. N Engl J Med 1982;**307**:1310-4.

11. Westert GP, Groenewegen PP. Medical practice variations: changing the theoretical approach. Scand J Public Health 1999;**27**:173-80.

12. Bisset AF. Appendicectomy in Scotland: a 20-year epidemiological comparison. J.Public Health Med. 1997;**19**:213-8.

13. Mulley AG. Inconvenient truths about supplier induced demand and unwarranted variation in medical practice. BMJ 2009;**339**:b4073.

14. Lilford RJ. Should the NHS strive to eradicate all unexplained variation? No. BMJ 2009;**339**:b4809.

15. Clarke MG, Wilson JR, Kennedy KP, MacDonagh RP. Clinical judgment analysis of the parameters used by consultant urologists in the management of prostate cancer. J Urol 2007;**178**:98-102.

16. Richards S. Should the NHS strive to eradicate all unexplained variation? Yes. BMJ 2009;**339**:b4811.

17. Sheldon TA, Cullum N, Dawson D, Lankshear A, Lowson K, Watt I et al. What's the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients' notes, and interviews. BMJ 2004;**329**:999.

18. Cookson R, McDaid D, Maynard A. Wrong SIGN, NICE mess: is national guidance distorting allocation of resources? BMJ 2001;**323**:743-5.

19. Pearson S, Littlejohns P. Reallocating resources: how should the National Institute for Health and Clinical Excellence guide disinvestment efforts in the National Health Service? J Health Serv Res Policy 2007;**12**:160-5.

20. Ball H, Kemp L, Fordham R. Road testing programme budgeting and marginal analysis: Norfolk Mental Health pilot project. Psychiatric Bulletin 2009;**33**:141-4.

21. Cook S. Disinvestment: Experts' guide to saving money in health. BMJ 2010;**340**:1281.

22. Leonard P. Shrinking budgets, improving care: Cut administration and buy wisely. BMJ 2010;**340**:1258.

23. Rich G. Shrinking budgets, improving care: Simplify processes and engage with patients. BMJ 2010;**340**:1251.

24. Zalmanovitch Y, Vashdi DR. Shrinking budgets, improving care: Trade-offs are unavoidable. BMJ 2010;**340**:1259.

25. Achieving world class productivity in the NHS 2009/10-2013/14. Obesity Surgery . 2009. McKinsey & Company.

26. Ibanez B, Librero J, Bernal-Delgado E, Peiro S, Lopez-Valcarcel BG, Martinez N et al. Is there much variation in variation? Revisiting statistics of small area variation in health services research. Bmc Health Services Research 2009;**9**.

27. Mindell J, Klodawski E, Fitzpatrick J, Malhotra N, McKee M, Sanderson C. The impact of private-sector provision on equitable utilisation of coronary revascularisation in London. Heart 2008;**94**:1008-11.

28. Birkmeyer JD, Sharp SM, Finlayson SRG, Fisher ES, Wennberg JE. Variation profiles of common surgical procedures. Surgery 1998;**124**:917-23.

29. Widmer M, Matter P, Staub L, Schoeni-Affolter F, Busato A. Regional variation in orthopedic surgery in Switzerland. Health & Place 2009;**15**:791-8.

30. Coory M, Gibberd R. New measures for reporting the magnitude of small-area variation in rates. Statistics in Medicine 1998;**17**:2625-34.

31. Martuzzi M, Hills M. Estimating the Degree of Heterogeneity Between Event Rates Using Likelihood. American Journal of Epidemiology 1995;**141**:369-74.

32. Goldstein H, Spiegelhalter DJ. League tables and their limitations: Statistical issues in comparisons of institutional performance. Journal of the Royal Statistical Society Series A-Statistics in Society 1996;**159**:385-409.

33. Marshall EG, Spiegelhalter DJ. Reliability of league tables of in vitro fertilisation clinics: retrospective analysis of live birth rates. British Medical Journal 1998;**316**:1701-4.

34. Greenland S, Robins JM. Empirical-Bayes Adjustments for Multiple Comparisons Are Sometimes Useful. Epidemiology /7;**2**:244-51.

35. Louis TA. Assessing, Accommodating, and Interpreting the Influences of Heterogeneity. Environmental Health Perspectives 1991;**90**:215-22.

36. Gelman A. Prior distributions for variance parameters in hierarchical models(Comment on an Article by Browne and Draper). Bayesian Analysis 2006;**1**:515-33.

37. Jones HE, Spiegelhalter DJ. The identification of 'unusual' health-care providers from a hierarchical model. (Submitted) 2010.

38. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement.Sci 2010;**5**:56.

39. Glaser BG, Strauss AL, Strutzel E. The Discovery of Grounded Theory; Strategies for Qualitative Research. Nursing Research 1968;**17**.

40. Heritage J. Conversation analysis and institutional talk. ed. Qualitative Research. Thousand Oaks: Sage; 2004.

41. Strauss A, Corbin JM. Basics of qualitative research: Grounded theory procedures and techniques. Thousand Oaks, CA, US: Sage Publications, Inc; 1990.

This protocol refers to independent research commissioned by the National Institute for Health Research (NIHR). Any views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the SDO programme or the Department of Health.