The clinical effectiveness and cost-effectiveness of telephone triage for managing same-day consultation requests in general practice: a cluster randomised controlled trial comparing general practitioner-led and nurse-led management systems with usual care (the ESTEEM trial)

John L Campbell,[1]* Emily Fletcher,[1] Nicky Britten,[2] Colin Green,[2] Tim Holt,[3] Valerie Lattimer,[4] David A Richards,[2] Suzanne H Richards,[1] Chris Salisbury,[5] Rod S Taylor,[1] Raff Calitri,[1] Vicky Bowyer,[6] Katherine Chaplin,[2] Rebecca Kandiyali,[1] Jamie Murdoch,[4] Linnie Price,[1] Julia Roscoe,[6] Anna Varley[4] and Fiona C Warren[1]

[1]Primary Care Research Group, University of Exeter Medical School, Exeter, UK
[2]Institute of Health Service Research, University of Exeter Medical School, Exeter, UK
[3]Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
[4]School of Nursing Sciences, Faculty of Medicine and Health Sciences, University of East Anglia, Norwich, UK
[5]Centre for Academic Primary Care, School of Social and Community Medicine, University of Bristol, Bristol, UK
[6]Department of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK

*Corresponding author

Declared competing interests of authors: none

Published February 2015
DOI: 10.3310/hta19130

Scientific summary

The effectiveness of telephone triage for same-day consultation requests

Health Technology Assessment 2015; Vol. 19: No. 13
DOI: 10.3310/hta19130

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

Demands on UK primary care are increasing, prompting an exploration of alternative ways of managing patients in an attempt to respond to government and societal expectations while continuing to deliver safe, high-quality care.

When combined with telephone consultation, it has been proposed that telephone triage improves the management of demand for primary care by providing more rapid access to health-care advice for patients, reducing non-attendance rates and freeing up opportunities for face-to-face consultation.

The majority of previous research relates to models of triage that involve nurses, with little research addressing the impact of general practitioner (GP) telephone triage. To date, there has been no large-scale randomised controlled trial (RCT) comparing the potential benefits and harms of GP- or nurse-led telephone triage (NT) of patients requesting same-day consultations.

Objectives

The overarching aim of this trial was, in comparison with usual care (UC), to assess the impact of NT- and GP-led telephone triage (GPT) on primary care workload and cost, patient experience of care, and patient safety and health status, for patients requesting same-day consultations in general practice. Specific objectives were to compare the effects on primary care workload and cost, and patient experience of care, patient safety and health status, of (1) NT vs. UC; (2) GPT vs. UC; and (3) NT vs. GPT. We also explored the experiences and views of patients and members of practice staff on the acceptability of telephone triage.

Methods

Design

Pragmatic cluster RCT incorporating economic evaluation and parallel qualitative process evaluation. A preliminary pilot RCT was conducted in six practices to (1) confirm the implementation of GP- and nurse-led triage systems as feasible; (2) confirm the proposed recruitment of practices and refine data collection systems; and (3) confirm the assumed level of clustering of outcomes.

Setting and participants

Forty-two general practices from four regions of England, UK (Devon, Bristol/Somerset, Warwickshire/Coventry and Norfolk/Suffolk). Participants were consecutive patients (aged $\geq$ 16 years or $<$ 12 years) seeking a same-day face-to-face consultation with a GP. Patients aged 12.0–15.9 years were excluded owing to concerns regarding confidentiality of a mailed questionnaire, as were patients with health-care needs that were deemed too urgent to wait for triage (e.g. difficulties breathing, chest pain) and patients who were unable to communicate in English by telephone.

Randomisation

Individual patient-level randomisation was deemed impractical as it does not reflect the practice-wide reality of triage system implementation and is vulnerable to contamination. Consenting practices (clusters) were randomised in a 1 : 1 : 1 ratio using a secure remote automated allocation system designed by a statistician who was independent of the research team. The allocation sequence was computer generated, and minimised for geographical location, practice deprivation and practice list size.

Interventions

Patients from practices allocated to either of the triage arms who requested a same-day face-to-face consultation with a GP were advised that they would be called by the clinician (GP or nurse using computer decision support software) later that day to discuss their needs and to discuss the most appropriate management option. The options available included (1) self-care advice; (2) a further within-practice consultation on the same day; (3) a further within-practice consultation on a subsequent day; and (4) referral to another NHS service. Triage practices received standardised training on how to implement triage, although some tailoring of the intervention was permitted to suit local needs. Patients in UC practices were managed following the standard protocols for that practice.

Outcome measures

Our primary outcome measure (POM) is the total number of primary care contacts occurring within a 28-day period following a patient's same-day consultation request. We defined primary care as including consultations within general practice (including the triage contact), walk-in centres, out-of-hours primary care services and attendance at accident and emergency (A&E). All POM data were collected by trained researchers conducting case note reviews of electronic practice records.

Other process and outcome data were collected to document NHS resource use, appointment non-attendance and consultation length. Secondary outcomes of patient safety (number of deaths and emergency hospital admissions within 7 days of the index request and A&E attendance at follow-up), NHS resource use and costs, appointment non-attendance and consultation length were collected from case note review and primary care Clinician Forms. Patient-reported health status [European Quality of Life-5 Dimensions (EQ-5D)] and experiences of care (modified English national GP Patient Survey items) were collected by postal questionnaire.

Given the nature of the interventions, it was not possible to blind patients or practitioners to treatment allocation. Although the cluster design of the trial might theoretically allow researchers who are undertaking case note reviews to be blinded, our pilot study showed this was not possible in practice. The data analysis was carried out by a statistician who was blinded to treatment allocation.

Sample size

Based on a previous UK study comparing NT with UC for handling same-day consultation requests, we powered our study to detect a between-arm difference of 1.02 [standard deviation (SD) 0.78] vs. 1.38 (SD 1.79) at follow-up between the triage arms and UC [based on 90% power, intracluster correlation coefficient (ICC) 0.05, two-sided alpha 0.05]. Assuming 47% non-collection of the POM, it was necessary to recruit 7046 patients from 14 practices across each arm to reach our target of 3751 patients per arm for analysis (i.e. a total of 21,138 patients from 42 practices). The pilot study provided confirmation of our assumed ICC of 0.05 (i.e. 0.03, 95% CI 0.00 to 0.08). Furthermore, the pilot study led to a change in the method of patient consent to case note review from written consent (obtained from completed patient questionnaires) to include initial verbal consent obtained from the treating clinician.

Statistical methods

The primary analysis of the POM took the form of a regression analysis using a hierarchical model to take account of the cluster allocation, using a random effect to adjust for potential clustering effect by practice, and allowing for adjustment for practice-level minimisation variables and patient-level covariates shown to differ at baseline. Models were performed twice, initially using the UC arm as reference, and then using the GPT arm as reference, to derive comparison between the two triage arms. Investigation of the effect of missing POM data (owing to lack of availability of a case note review) was undertaken using multiple imputation methods, based on the assumption that missing case note review data were missing at random. Additional analyses were conducted on the POM, and on secondary measures derived from the POM, using the hierarchical generalised linear model methods described above. Some of these analyses were determined a priori; others were determined post hoc following initial inspection of the data.

Economic methods

Primary economic (cost) analyses were undertaken using data collected on the POM contacts taking place over 28 days, and conducted from the perspective of the NHS. Economic analyses were conducted in accordance with the statistical methods above. The primary economic analysis estimated the mean cost of care across each of the trial arms, to include triage (where used) and the items in the primary outcome. Analysis was based on a microlevel costing estimate for the triage intervention, and the use of published unit cost data for other elements of resource use. Estimates of the cost associated with triage interventions were based on incremental costs when compared with UC, with any capital costs and/or training costs depreciated/spread over an appropriate time period in the primary analyses (with other time horizons for these costs explored in sensitivity analyses). Sensitivity analyses were undertaken against the primary analyses to explore the implications of uncertainty in data used and the assumptions made within the primary analyses.

Results

Participating practices ($n = 42$) and patients were well balanced across the three trial arms with respect to key characteristics (practice list size, setting and deprivation; patient age, gender and deprivation). In total, 20,990 patients were eligible for the trial (UC $n = 7283$; GPT $n = 6695$; NT $n = 7012$). POM data were analysed for 16,211/20,990 (77%) participants (UC $n = 5572$; GPT $n = 5171$; NT $n = 5468$). There was some evidence of participation bias, with young adults (aged 16–24 years) being less likely than the reference age group (25–59 years), and women being less likely than men, to have POM data available.

The mean number of POM contacts in the 28-day follow-up period was 1.91 in UC (SD 1.43; total 10,616), 2.65 in GPT (SD 1.74; total 13,720) and 2.81 in NT (SD 1.68; total 15,400). Compared with UC, there was an increase in POM contacts of 33% in GPT [rate ratio (RR) 1.33, 95% CI 1.30 to 1.36] and 48% in NT (RR 1.48, 95% CI 1.44 to 1.52). There was a small increase of 4% in NT (RR 1.04, CI 1.01 to 1.08) compared with GPT.

Triage-arm patients had more diverse patterns of management when compared with UC. For GP face-to-face and telephone contacts combined across the 28-day follow-up period, the RR was 1.38 (95% CI 1.28 to 1.50) in GPT compared with UC, and 0.84 (95% CI 0.78 to 0.91) in NT compared with UC. GP face-to-face contacts decreased by 39% during the 28-day follow-up period in GPT compared with UC (RR 0.61, 95% CI 0.54 to 0.69) and in NT by 20% compared with UC (RR 0.80, 95% CI 0.71 to 0.90). Following the implementation of triage, no impact was observed on contacts with other services outside of the practice (out-of-hours primary care, walk-in centres or A&E) in either intervention arm.

Changes were also observed in the distribution of estimated patient–clinician contact time on the index day following the introduction of triage. Introducing GPT was associated with a small increase in patient–nurse contact time, whereas introducing NT was associated with both a decrease in patient–GP contact time and a substantial increase in overall patient–nurse contact time.

The estimated health-care costs over the 28-day follow-up were similar across all three arms, at a mean cost of approximately £75 (US$120, €88) per patient.

There was no evidence of differences across the three trial arms with respect to patient safety (patient mortality, emergency hospital admissions and A&E attendance rates) and patient health status did not vary between triage arms and UC, or between GPT compared with NT. NT was somewhat less acceptable to patients than GPT or UC.

Data from 84 qualitative interviews with patients and staff (sampled from 10 practices) found no strong, compelling or consistent narrative about what works and what does not work when implementing telephone triage in primary care. Rather, the qualitative data highlighted the complexity of primary care organisations and the significance of individual practice culture. Both triage models were sometimes experienced positively by staff and patients, whereas others viewed it negatively.

Conclusions

We believe ESTEEM to be the first RCT to assess both nurse- and GP-led telephone triage of patients requesting same-day consultations in primary care. The ESTEEM trial achieved its recruitment target of 22,000 patients across 42 general practices in four regions in England.

Contrary to suggestions that triage can reduce primary care workload, we found that both GPT and NT increased the number of primary care contacts in the 28 days following a patient's same-day consultation request compared with maintaining a UC approach to such requests. Following the introduction of GPT, we observed a redistribution of GP workload from face-to-face to telephone consultations. Following the introduction of NT, we observed a redistribution of workload from GPs to nurses. Triaged patients had more diverse patterns of management than those in UC, possibly indicating more flexible approaches to patient management in the triage arms. There was evidence of increased nurse deployment in both triage arms, with this being substantial following the introduction of NT.

When considering the differing patterns and duration of patient care contacts, both forms of triage were cost neutral to the NHS compared with UC. However, we found no important gains in patients' safety, health status or experience of care with triage.

Our process evaluation identified no strong, compelling or consistent narrative about what works and what does not work when implementing telephone triage in primary care. It did, however, identify key issues on which any practice considering implementing triage might want to reflect before adopting a triage system. These included consideration of issues relating to individual practice culture and capacity, and the forward planning, with the whole staff team, of any major changes in access arrangements.

Our results do not support a definitive policy recommendation to roll out across the NHS either nurse or GP triage for the management of same-day appointments. Triage, whether implemented by a GP or by a nurse using decision support software, should be introduced with full awareness of the whole-system implications arising from the decision to implement such a process. Notwithstanding this, clinician triage of patients seeking same-day consultations may offer advantages in supporting the flexible delivery of patient care, and potentially offers a useful approach in the armamentarium of tools facilitating the delivery of effective NHS primary care.

Trial registration

This trial is registered as ISRCTN20687662.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 5.116

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 08/53/15. The contractual start date was in November 2009. The draft report began editorial review in November 2013 and was accepted for publication in April 2014. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).