Evaluation of appropriateness methods to define and improve access across primary, secondary and tertiary care among people with angina

Revised report to the National Co-ordinating Centre for NHS Service Delivery and Organisation R&D (NCCSDO)

May 2006

prepared by:
Professor Harry Hemingway¹
Professor Gene Feder²
Dr Cornelia Junghans¹
Dr Claire Somerville²,
Dr Katie Featherstone³
Natalie Fitzpatrick¹
Dr Sarah Cotter²
Professor Nick Black⁴
Dr Paul Shekelle⁵

With the help of:
Dr Angela Crook¹
Dr Ruoling Chen²
Dr Neha Sekhri⁶,
Dr Melvyn Jones⁷
Professor Sheila Hillier⁸,
**Affiliations**

1. Department of Epidemiology and Public Health, University College London Medical School.
2. Centre for Health Sciences, Queen Mary’s School of Medicine and Dentistry
3. School of Nursing and Midwifery Studies, Cardiff University
4. Department of Public Health and Policy Health Services Research Unit, London School of Hygiene and Tropical Medicine
5. Greater Los Angeles Veterans Affairs Health Care System, Los Angeles, California
6. Barts and the London, Queen Mary's School of Medicine and Dentistry
7. Department of Primary Care and Population Sciences, University College London Medical School
8. Institute of Health Sciences Education, Queen Mary's School of Medicine and Dentistry

We gratefully acknowledge the work of Dr Sarah Cotter in undertaking the analyses for this report.
Executive summary

This study was commissioned by the NHS Service and Delivery Organisation R&D programme. It examined the effectiveness and feasibility of appropriateness ratings devised by experts in improving appropriate and equitable access to investigations in patients with angina.

Access to investigation (i.e. tests) for angina remains highly variable; coronary mortality among patients who are not investigated is higher than in the general population. Conventional guideline recommendations such as those endorsed by NICE tend to relate to broad groups of patients rather than individual, specific patient characteristics, and such conventional guidance has had limited success in reducing practice variations or altering clinician behaviour.

Appropriateness ratings constitute a novel decision support technology, which incorporates expert panel recommendations for individualised, specific patient indications. The impact of such specific decision support on physician decision making has been unclear. Also unclear is the acceptability and feasibility of such ratings in clinical practice and whether factors which are not incorporated into such ratings, for example patient preference, non-attendance rates or other patient or clinician characteristics would influence investigation and treatment despite the use of appropriateness technology.

Given the complex nature of the enquiry we used several data sources and analysis techniques in an attempt to answer the following three questions and meet the four associated research objectives.

**Question 1: How effective are appropriateness ratings compared with conventional guidelines in improving appropriate and equitable decision making in primary and secondary care physicians?**

- **Objective 1** To compare the effect of specific vs. conventional guidelines on physician recommendations for investigation of angina as a means to improving appropriate access

- **Objective 2** To determine impact on access to investigation among older patients, women and south Asians. (*ARIA randomised controlled trial*)

**Question 2: How acceptable are appropriateness ratings to clinicians and patients and what barriers to their use exist in day to day clinical practice?**

- **Objective 3** To determine the acceptability of appropriateness ratings to clinicians and patients. (*ARIA-QUALitative research*)
Question 3: What other barriers may exist further down the patient pathway that may prevent appropriate care despite the use of appropriateness ratings?

Objective 4: To determine the extent of physician and patient barriers in mediating underuse of revascularization. (ACRE secondary data analysis).

Methods and Results

ARIA randomised controlled trial (Objectives 1 and 2)

Interventions

One trial arm received decision support via appropriateness ratings individualised to specific patient characteristics, which were rated by two independent expert panels in agreement, the other trial arm received decision support via conventional guidelines from the American Heart Association, European Society of Cardiology and North of England.

Participants and patient vignettes

292 clinicians (189 general practitioners and 103 cardiologists) from 9 regions in the UK were recruited. 145 were randomised to patient-specific appropriateness ratings and 147 to conventional guidelines. Clinicians were unaware of the trial objective. Each clinician made recommendations on investigation of 24 patient vignettes (from a pool of 48 unique vignettes): 12 without and then 12 with decision support intervention. We analysed 5938 decisions for exercise ECG and 6291 decisions for angiography.

Outcomes

The primary outcome was the proportion of recommendations made by doctors that agreed with those made by two independent national expert panels as well as the proportion of recommendations made by clinicians in agreement with their own intervention. As a secondary outcome we examined the effect of appropriateness ratings and guidelines on appropriate decision making in pre-specified patient groups (by age, sex and ethnicity).

ARIA – QUAL design and participants (Objective 3)

Interviews with a purposeful sample of 14 patients attending a rapid access or a cardiology clinic. Interviews with doctors (20), ECG technicians (10) and administrators (2) based in the clinics. Observation of 115 clinical consultations and interviews with doctors (20), patients (14), ECG technicians (10) and administrative staff (2). Interviews on use of decision support tool with doctors (8), technicians (4) and a research manager (1). The everyday work of one clinical cardiology team (10 doctors and 8 ECG technicians over the fieldwork period) was observed over fourteen weeks. We observed 115
individual consultations in the two clinics. Six patients (5 men and 1 woman) recruited from the cardiac rehabilitation clinic participated in an extended (3 hours) focus group.

Data sources

Interviews with six doctors who piloted the ARIA decision support tool using data from 15 consultations that they had just completed. Interviews with 10 doctors who participated in the ARIA trial. Focus group of six patients recruited from the cardiac rehabilitation clinic.

Analysis

Based on transcripts of audio-tape recorded interviews and transcripts of the detailed notes taken during and immediately following periods of observation. Thematic analysis using grounded theory significant themes underpinning the process of diagnosis and decision-making, from patient and clinician perspectives.

ACRE design and participants (Objective 4)

We abstracted case note data among participants from the original ACRE cohort who consented to 7 year follow-up (n=2593) in order to select and compare management after angiography in patients who were rated appropriate for CABG by the ACRE expert panel (n=927). We compared patients in whom CABG was rated appropriate who did (n= 530) and did not (n = 397) undergo CABG in terms of important patient, clinical and care characteristics. We used logistic regression for analysis.

Results

Our work confirmed that conventional guidelines had little effect on outcome within or between trial arms for either investigation, or in either speciality. After intervention, the proportion of recommendations made in agreement with the expert panel recommendations for exercise ECG increased (819/1281 (63%) vs. 619/1281 (48%), P<0.0001 and for coronary angiography 1274/2292 (56%) vs. 1018/2292 (44%) P<0.0001 with the ratings arm. Within trial arm comparisons, using clinicians as their own controls, showed very similar effects. Pre-defined analyses of sub-groups for which inequitable access has been observed, showed that improvements in agreement with specific recommendations were robust at older ages, among women, and among south Asians. The change in decisions associated with specific decision support resulted in an increased use of exercise ECG and angiography for both GPs and cardiologists.

While the trial demonstrated a clear proof of concept and showed that appropriateness ratings have significant potential to improve clinical decision making, our findings from the qualitative work show that any implementation of appropriateness ratings needs to consider the clinical culture and the
nature of clinical decision making. We identified a variety of criteria which were influential in unaided decision making in the diagnosis of angina, such as patient perspective (development of the angina/heart disease narrative, role of family or work stressors, ambiguity of family history, faith in cardiological interventions and notions of invincibility), clinical context and role of place, including the diversity of consultations, the effect of the physical clinic space and its impact on decision making and systemic factors.

In terms of acceptability and feasibility of appropriateness ratings in clinical practice we found that several reasons accounted for ordering an inappropriate investigation, for example patient or doctor reassurance and information emerging in the consultation that is not a variable used in the ARIA tool (e.g. inability to perform ETT or inconclusive result when deciding about an angiogram), language barrier making history less reliable and therefore lowering threshold for an ETT. Reasons for underuse of appropriate investigations included conviction of a diagnosis of non-cardiac chest pain or patient’s unfitness for ETT or angiogram due to obesity or co-morbidity for example. We found scepticism of junior and senior doctors working in the cardiac team contrasted with relative acceptance by GP informants regarding the concept of an expert panel. The concept of expertise in relation to formal recommendations embedded in decision support tool was seen as problematic by many informants. But conventional guidelines were perceived as even less reliable. Focus group patients were generally enthusiastic about the notion of decision support technology to aid clinical decision making.

In summary, the dialogue between patient and clinician leading up to a working diagnosis and the decision to investigate further were highly complex with multiple influences. The use of the decision support tool raised issues about the face validity of experts and about the role of patient factors that were not included in the decision support variables. The work highlighted the fact that tests were often carried out for reasons independent of appropriateness, such as reassurance of patient and doctor, a means of communicating with other doctors or hesitance to disagree with senior doctors. Interestingly, patients attending chest pain clinics were sometimes unaware of the nature of the clinic and ignorant of the fact that they were seeing a heart specialist.

On a large scale, patient preference, non-attendance or clinician outliers were not found to explain the significant underuse of revascularization in appropriate patients. For CABG this underuse ranged between clinicians from 54-75%, for PCI it ranged from 36-78%. Patient preference or non-attendance was virtually unrecorded in a large sample of case notes. This raises important questions about the way patient involvement and decisions of the clinical team are documented. It also calls for further explorations of why this substantial underuse of revascularization occurs, as often cited reasons such as patient preference or clinician outliers do not seem to explain this phenomenon. Our findings from the ACRE analysis as well as ARIA-QUAL suggest that a whole system approach is needed as it is likely to be the small variation within all clinicians who make investigation decisions often for
Appropriateness methods for defining and improving access to angina care

reasons other than clinical appropriateness, rather than a few doctors making consistently inappropriate decisions.

**Overall conclusion**

Specific decision support lead to changes in clinician recommendations for the investigation of angina, but widely advocated conventional guidelines had very little effect despite tipping the scales in their favour by providing easily accessible electronic presentation of the relevant guideline sections. This effect was robust among vignettes in older people, women and south Asians. Randomised trials in real patients should now be carried out to test whether specific decision support can improve access to care and, ultimately, patient prognosis. Based on the findings from ARIA-QUAL such a trial should be designed as a complex intervention and need to address shortcomings of the appropriateness method and reasons why clinicians do not follow the recommendations. Patient preferences, non-attendance rates and undue physician variation seem to play a minor role in influencing investigation in appropriate patients, although it is inconclusive whether this is due to lack of reporting on clinical notes or true absence of these factors.

**Implications for clinical practice, policy and access and disparity research**

Our findings support the consideration of:

- Randomised trials of appropriateness tools in decision support in real patients, with assessment of clinical outcomes, which should be conceived and developed as a complex intervention.

- Expert panel ratings need to be further developed and improved, so that they have higher reliability (less disagreement amongst experts) and deal with a rapidly changing clinical environment, lack of evidence and missing data in real patients.

- The need for a whole system approach to designing interventions to improve access to care, as highlighted by our finding that variation was not due to some clinicians being outliers but due to consistent underuse of revascularization by all clinicians.

- Patient involvement in decision making. A prospective study in which consultations are observed might elucidate the role of patient preference. Our findings show that despite increased patient involvement preferences rarely influence clinical decisions.

- Better primary studies on which to base investigation decisions (for example randomised trials of different investigation strategies) are needed. Appropriateness ratings may help identify focussed areas of uncertainty where evidence is needed.
Disclaimer

This report presents independent research commissioned by the National Institute for Health Research (NIHR). The 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.

Addendum

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 and 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