Open Access: free full text available online
For the full versions of these articles see bmj.com

Incremental prognostic value of the exercise electrocardiogram in the initial assessment of patients with suspected angina: cohort study

Neha Sekhri,[1] Gene S Feder,[4] Cornelia Junghans,[3] Sandra Eldridge,[2] Athavan Umaipalan,[2] Rashmi Madhu,[2] Harry Hemingway,[3] Adam D Timmis[2]

EDITORIAL by Abramson

[1]Newham University Hospital, London
[2]Barts and The London Queen Mary's School of Medicine and Dentistry, London
[3]Department of Epidemiology and Public Health, University College London Medical School, London
[4]Unit of Academic Primary Health Care, University of Bristol
Correspondence to: A D Timmis, London Chest Hospital, Barts and The London NHS Trust, London E2 9JX
adamtimmis@mac.com

Cite this as: *BMJ* 2008;337:a2240
doi:10.1136/bmj.a2240

ABSTRACT

Objective To determine whether resting and exercise electrocardiograms (ECGs) provide prognostic value that is incremental to that obtained from the clinical history in ambulatory patients with suspected angina attending chest pain clinics.

Design Multicentre cohort study.

Setting Rapid access chest pain clinics of six hospitals in England.

Participants 8176 consecutive patients with suspected angina and no previous diagnosis of coronary artery disease, all of whom had a resting ECG recorded. 4848 patients with a summary exercise ECG result recorded (positive, negative, equivocal for ischaemia) comprised the summary ECG subset of whom 1422 with more detailed exercise ECG data recorded comprised the detailed ECG subset.

Main outcome measure Composite of death due to coronary heart disease or non-fatal acute coronary syndrome during median follow-up of 2.46 years.

Results Receiver operating characteristics curves for the basic clinical assessment model alone and with the results of resting ECGs were superimposed with little difference in the C statistic. With the exercise ECGs the C statistic in the summary ECG subset increased from 0.70 (95% confidence interval 0.68 to 0.73) to 0.74 (0.71 to 0.76) and in the detailed ECG subset from 0.74 (0.70 to 0.79) to 0.78 (0.74 to 0.82). However, risk stratified cumulative probabilities of the primary end point at one year and six years for all three prognostic indices (clinical assessment only; clinical assessment plus resting ECG; clinical assessment plus resting ECG plus exercise ECG) showed only small differences at all time points and at all levels of risk.

Conclusion In ambulatory patients with suspected angina, basic clinical assessment encompasses nearly all the prognostic value of resting ECGs and most of the prognostic value of exercise ECGs. The limited incremental value of these widely applied tests emphasises the need for more effective methods of risk stratification in this group of patients.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2240

INTRODUCTION

The character of symptoms in patients with suspected angina is generally accepted to be central to clinical diagnosis and risk assessment.[1-3] Usually a resting electrocardiogram (ECG) is recorded and often an exercise ECG, which is a routine investigation in 59% of the chest pain clinics in the United Kingdom.[4] Whether these investigations contribute incrementally to risk assessment in this group is unknown.

We studied a large cohort of ambulatory patients admitted to rapid access chest pain clinics with recent onset of suspected angina. None had been previously assessed for cardiovascular disease. We identified clinical predictors of coronary events and determined whether resting and exercise ECGs provide prognostic value incremental to that obtained from the clinical history.

METHODS

Overall, 8176 of 10 634 consecutive patients with new onset of chest pain referred to six chest pain clinics from 1 January 1996 to 31 December 2002 were included in this study (see bmj.com). These patients comprised the cohort; 4873 (60%) with an exercise ECG recorded were then stratified into two subsets—those with summary test results (n=4848) and those with additional detailed test data (n=1422). The exercise ECG was obtained in all but 7% of patients.

We recorded clinical data in a customised database[5] at the time of the consultation (see bmj.com). After the consultation we recorded the diagnosis based on the clinical assessment (angina, non-cardiac chest pain, other). We obtained a resting 12 lead ECG for every patient, recorded as normal or abnormal according to rhythm, conduction, change in ST segment or T wave, left ventricular hypertrophy, and Q waves. Treadmill stress testing was undertaken according to perceived clinical need. In the summary ECG subset only the clinicians' assessment of ischaemia was recorded (positive, negative, or equivocal). In the detailed ECG subset, data recorded included exercise time; maximum workload, heart rate, and blood pressure; change in ST segment; arrhythmias; and reason for stopping.

We used the primary discharge diagnosis to define events among patients undergoing hospital admission during the follow-up period. The primary end point was a

composite of death due to coronary heart disease or non-fatal acute coronary syndrome.[6]

Statistical analysis

We used Stata version 8.0 for analyses. We compared the characteristics of the patients in the cohort with those in the subsets. Data on patients who underwent coronary artery bypass grafting or percutaneous coronary intervention were censored at the time of the revascularisation procedure. Firstly, we carried out multivariable Cox analysis for the primary end point using factors that were significant at the 20% level in univariable analysis for each of three separate models: clinical (age, sex, typicality of symptoms, heart rate, systolic blood pressure, history of hypertension, diabetes, smoking status), ECG (QRS axis deviation, Q waves, change in ST segment or T wave, left ventricular hypertrophy, bundle branch block), summary exercise ECG (positive, negative, or equivocal), and detailed exercise ECG (exercise time, maximum workload, percentage predicted heart rate, maximum blood pressure, reason for stopping, change in ST segment, exertional arrhythmias). We then used the covariates that remained significant at the 5% level in each model to build three incremental models: clinical assessment, clinical assessment plus resting ECG, and clinical assessment plus resting ECG plus either summary or detailed exercise ECGs. Then we calculated prognostic indices for each of these models using the regression coefficients. To maximise power we developed the basic clinical model and the resting ECG model in the whole cohort and used the subsets to develop the models that include data on the exercise ECG. We forced the variables from the basic clinical model and the resting ECG model into the models with data on the exercise ECGs. We plotted receiver operating characteristics curves and calculated the C statistic with confidence intervals for each prognostic index[7] to examine the incremental prognostic value of the resting and exercise ECGs. Curves were plotted for all patients and for patients with intermediate probability of coronary artery disease, based on an algorithm that takes into account age, sex, and typicality of chest pain.[8] Finally, we arbitrarily split the prognostic index into thirds for risk[9] and charted the probabilities of developing an event at one year and six years to illustrate the incremental value of the resting ECG and exercise ECG for each risk category during long term follow-up.

RESULTS

Participants in the cohort were followed-up for a median 2.46 years (interquartile range 1.61-3.92 years), in the summary ECG subset for 2.21 (1.27-3.26), and in the detailed ECG subset for 2.26 (1.51-5.18). Participants in the subsets had similar characteristics to those in the cohort except for a greater proportion of patients with atypical symptoms and an intermediate probability of coronary artery disease (see bmj.com). Angina was diagnosed in 29% of the cohort, 32% of the summary ECG subset, and 28% of the detailed ECG subset.

Covariate screening and prognostic indices

In the clinical assessment model the variables of typical chest pain, age, diabetes, and being male were independently associated with an increased risk of the primary end point (see bmj.com). In the resting ECG model, variables independently associated with an increased risk of the primary end point were bundle branch block, change in ST segment or T wave, and Q waves. In the summary exercise ECG model the result was associated with the primary end point, and in the detailed exercise ECG model exercise time and change in the ST segment on exertion were independently associated.

In the final iterations (clinical assessment plus resting ECG plus exercise ECG) the major contributors to the risk of the primary end point were typical symptoms and abnormalities on the exercise ECG, with age, sex, and diabetes making variable contributions depending on subset analysed.

Cumulative probabilities of an event according to three prognostic indices in cohort and in exercise electrocardiogram (ECG) subsets

	% cumulative probability (95% CI) at 1 year			% cumulative probability (95% CI) at 6 years		
Risk groups	Clinical assessment	Clinical assessment plus resting ECG	Clinical assessment plus resting ECG plus exercise ECG*	Clinical assessment	Clinical assessment plus resting ECG	Clinical assessment plus resting ECG plus exercise ECG*
Cohort:						
Low	1 (0 to 1)	1 (0 to 1)	NA	4 (3 to 6)	4 (3 to 6)	NA
Medium	2 (2 to 3)	1 (1 to 2)	NA	10 (8 to 12)	8 (7 to 11)	NA
High	7 (6 to 8)	8 (7 to 9)	NA	25 (22 to 28)	27 (23 to 30)	NA
Summary ECG subset:						
Low	1 (1 to 2)	0 (1 to 2)	1 (1 to 2)	5 (3 to 8)	5 (3 to 8)	5 (3 to 8)
Medium	2 (2 to 3)	3 (2 to 3)	3 (2 to 3)	11 (9 to 15)	11 (8 to 14)	11 (8 to 14)
High	8 (6 to 9)	8 (6 to 9)	8 (6 to 9)	23 (20 to 27)	24 (21 to 29)	24 (21 to 27)
Detailed ECG subset:						
Low	0 (0 to 2)	0 (0 to 1)	0 (0 to 1)	5 (3 to 10)	3 (1 to 7)	3 (1 to 8)
Medium	2 (1 to 4)	2 (1 to 4)	3 (2 to 5)	8 (5 to 13)	9 (6 to 14)	9 (6 to 14)
High	8 (6 to 10)	8 (6 to 11)	7 (5 to 10)	21 (16 to 26)	22 (17 to 28)	22 (17 to 28)

NA=not applicable.
*Either incremental summary exercise ECG or detailed exercise ECG depending on subset.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Accurate identification of patients with suspected angina at higher risk of acute coronary syndromes and death is essential to tailor management strategies

Resting and exercise electrocardiography are the most widely performed non-invasive tests in patients with suspected angina but their contribution to risk assessment is unknown

Most research has focused on newer and more costly tests for diagnosing coronary disease, with less attention to incremental value for risk assessment

WHAT THIS STUDY ADDS

Clinical assessment of patients with suspected angina embraces most of the prognostic information provided by resting ECGs and by exercise ECGs

The incremental prognostic value of detailed exercise ECG variables is little better than the summary assessment of ischaemia commonly used

Receiver operating characteristics curves and C statistic

In the cohort, receiver operating characteristics curves for the clinical assessment model alone and with iteration for the resting ECG were superimposed with little or no increment in the C statistic (see bmj.com). With the iterations for the exercise ECGs the C statistic for the clinical assessment model increased in the summary ECG subset from 0.70 (95% confidence interval 0.68 to 0.73) to 0.74 (0.71 to 0.76) and in the detailed ECG subset from 0.74 (0.70 to 0.79) to 0.78 (0.74 to 0.82). When analysis was restricted to patients with an intermediate probability of coronary artery disease, the receiver operating characteristics curves for the clinical assessment model alone and with the ECG iteration remained superimposed, reflecting poor discrimination. With the exercise ECG iterations the C statistic (95% confidence interval) for the clinical assessment model increased in the summary ECG subset from 0.69 (0.65 to 0.73) to 0.74 (0.70 to 0.78) and in the detailed ECG subset from 0.69 (0.62 to 0.77) to 0.76 (0.70 to 0.82).

Patient outcomes

Typical chest pain and abnormalities on the resting and exercise ECGs were associated with adverse outcomes (see bmj.com). Thus point estimates of the probability of the primary end point at three years were 16% for typical chest pain, 15% for an abnormal resting ECG, and 19% for an abnormal exercise ECG, compared with 3%, 5%, and 9% for non-specific chest pain and normal resting and exercise ECGs. However, 47% (n=166) of the events occurred in patients with a "normal" exercise ECG. Thus in both subsets, risk stratified cumulative probabilities of the primary end point at one year and six years for all three prognostic indices (clinical, clinical plus resting ECG, clinical plus resting ECG plus exercise ECG) showed only small differences at all time points and in all thirds of risk (table).

DISCUSSION

In ambulatory patients with suspected angina, the clinical assessment embraces nearly all the prognostic information provided by the resting and exercise ECGs. The limited incremental value of these widely applied investigations extended across all thirds of risk, emphasising the importance of the clinical assessment and the need for more effective methods of risk stratification in this group of patients.

In patients with undifferentiated chest pain there is a gradient of coronary risk, greatest in those with a diagnosis of angina but extending to those with non-cardiac chest pain.[10] Our patients with typical angina were at higher risk of adverse outcomes than patients with atypical symptoms. Experience of non-invasive testing in patients with undifferentiated chest pain has increased in recent years, but uncertainty about its value for risk assessment remains.[11] We found that a range of abnormalities in both the resting ECG and the exercise ECG were independently predictive of adverse events in ambulatory patients with chest pain of recent onset. Overall, however, 47% of events during follow-up occurred in patients with a normal exercise ECG, emphasising the limitations of using ECGs for risk assessment. In this respect our findings were unequivocal, particularly for the resting ECG, which showed no incremental prognostic value above that of the clinical assessment. For the exercise ECGs, incremental prognostic value seemed greater, as the point estimates for the C statistic increased by 5.7% for the summary ECG subset and by 5.1% for the detailed ECG subset. Increases were only marginally greater among patients with an intermediate probability of coronary artery disease, the group in which the exercise ECG is most useful for diagnostic purposes.[12] In clinical terms the incremental prognostic value was trivial (see table), with the indices that incorporated data from the exercise ECG proving no more effective than those of the clinical assessment in predicting adverse outcomes. Detailed analysis of variables in the exercise ECG performed little better than the summary assessment commonly used in clinical practice.

Previous studies of treadmill testing in patients with chest pain have been small and have not been powered to test prognostic value.[11] Before these studies, a study[13] had developed a prognostic treadmill score in patients referred for cardiac catheterisation, which has since been validated in outpatient populations.[14] The incremental value of the score for predicting survival at four years was modest and similar to our exercise ECG model.[14]

Our methods are robust and reflect the incremental value of the statistical models rather than simple comparison of likelihood ratios. Only 7% of patients (n=725) were excluded. Outcomes in this group were not significantly different from those included in our study. Important limitations were the absence of data on lipid levels and family history, although they would probably have improved the discriminatory power of the clinical model. Similarly, our conclusions would not have been affected by exclusion of the 167 patients (2%) with Q waves, suggesting a history of silent infarction, because this would have further reduced the prognostic value of the resting ECG. An abnormal exercise ECG influences decisions on revascularisation, which in turn might influence prognosis. A sensitivity analysis using coronary artery bypass grafting as part of the composite end point did not change the findings. One of the most consistent

prognostic markers in exercise testing is maximum exercise capacity.[12] The prognostic importance of age is not undermined by its lack of significance in the final incremental model, which merely indicates that in patients undergoing exercise testing, more prognostic weight is contributed by exercise time, change in the ST segment, and typical chest pain.

Our study emphasises the importance of the clinical assessment for prognosis in patients with suspected angina. The data show that the need to improve risk stratification cannot be met by the resting ECG whereas the incremental value of the exercise ECG is small.

Contributors: See bmj.com.
Funding: The study was funded by the National Health Service's Service Delivery and Organisation research and development programme, to which interim progress reports were submitted. The funding body was not involved in the study design or analysis.
Competing interests: None declared.
Ethical approval: This study was approved by the multiregional ethics committee (/02/04/095). Permission was given by the National Patient Information Advisory Group (www.advisorybodies.doh.gov.uk/piag/) to link anonymised datasets without individual patient consent.
Provenance and peer review: Not commissioned; externally peer reviewed.

1 Diamond GA. A clinically relevant classification of chest discomfort. *J Am Coll Cardiol* 1983;1(2 Pt 1):574-5.
2 Pryor DB, Shaw L, McCants CB, Lee KL, Mark DB, Harrell FE Jr, et al. Value of the history and physical in identifying patients at increased risk for coronary artery disease. *Ann Intern Med* 1993;118:81-90.
3 Constant J. The clinical diagnosis of nonanginal chest pain: the differentiation of angina from nonanginal chest pain by history. *Clin Cardiol* 1983;6:11-6.
4 Timmis A, Sekhri N, Feder G, Hemingway H. Evaluation of rapid access chest pain clinics. 2007. www.sdo.lshtm.ac.uk/files/project/32-final-report.pdf.
5 Ray S, Archbold RA, Preston S, Ranjadayalan K, Suliman A, Timmis AD. Computer-generated correspondence for patients attending an open-access chest pain clinic. *J R Coll Physicians Lond* 1998;32:420-1.
6 ICD-10 for mortality. 2005. www.statistics.gov.uk/about/classifications/icd10/default.asp.
7 Newson R. Confidence intervals for rank statistics: percentile slopes, differences, and ratios. *Stata J* 2006;6:497-520.
8 Diamond GA, Forrester JS. Probability of CAD. *Circulation* 1982;65:641-2.
9 Rassi A Jr, Rassi A, Little WC, Xavier SS, Rassi SG, Rassi AG, et al. Development and validation of a risk score for predicting death in Chagas' heart disease. *N Engl J Med* 2006;355:799-808.
10 Sekhri N, Feder GS, Junghans C, Hemingway H, Timmis AD. How effective are rapid access chest pain clinics? Prognosis of incident angina and non-cardiac chest pain in 8762 consecutive patients. *Heart* 2007;93:458-63.
11 Gibbons RJ, Balady GJ, Bricker JT, Chaitman BR, Fletcher GF, Froelicher VF, et al. ACC/AHA 2002 guideline update for exercise testing: summary article. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1997 Exercise Testing Guidelines). *J Am Coll Cardiol* 2002;40:1531-40.
12 Gibbons RJ, Balady GJ, Beasley JW, Bricker JT, Duvernoy WF, Froelicher VF, et al. ACC/AHA guidelines for exercise testing. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Exercise Testing). *J Am Coll Cardiol* 1997;30:260-311.
13 Mark DB, Hlatky MA, Harrell FE Jr, Lee KL, Califf RM, Pryor DB. Exercise treadmill score for predicting prognosis in coronary artery disease. *Ann Intern Med* 1987;106:793-800.
14 Mark DB, Shaw L, Harrell FE Jr, Hlatky MA, Lee KL, Bengtson JR, et al. Prognostic value of a treadmill exercise score in outpatients with suspected coronary artery disease. *N Engl J Med* 1991;325:849-53.

Accepted: 10 September 2008

Efficacy of percutaneous versus intradermal BCG in the prevention of tuberculosis in South African infants: randomised trial

Anthony Hawkridge,[1] Mark Hatherill,[1] Francesca Little,[2] Margaret Ann Goetz,[3] Lew Barker,[3] Hassan Mahomed,[1] Jerald Sadoff,[3] Willem Hanekom,[1] Larry Geiter,[3] Greg Hussey,[1] the South African BCG trial team

EDITORIAL by von Reyn and Zumla

[1]South African Tuberculosis Vaccine Initiative, University of Cape Town, South Africa
[2]Department of Statistical Science, University of Cape Town
[3]Aeras Global TB Vaccine Foundation, Rockville, MD, USA
Correspondence to: A Hawkridge, Aeras Global TB Vaccine Foundation, Africa Office, Belmont Square, Rondebosch, Cape Town 7700 South Africa
thawkridge@aeras.org

Cite this as: *BMJ* 2008;337:a2052
doi:10.1136/bmj.a2052

ABSTRACT

Objective To compare the incidence of tuberculosis over two years in infants vaccinated at birth with intradermal BCG or with percutaneous BCG.

Design Randomised trial.

Setting South Africa.

Participants 11 680 newborn infants.

Interventions Infants were randomised by week of birth to receive Tokyo 172 BCG vaccine through the percutaneous route (n=5775) or intradermal route (n=5905) within 24 hours of birth and followed up for two years.

Main outcome measures The primary outcome measure was documented *Mycobacterium tuberculosis* infection or radiological and clinical evidence of tuberculosis disease. Secondary outcome measures were rates of adverse events, all cause and tuberculosis specific admissions to hospital, and mortality.

Results The difference in the cumulative incidence of definite, probable, and possible tuberculosis between the intradermal group and the percutaneous group, as defined using study definitions based on microbiological, radiological, and clinical findings was −0.36% (95.5% confidence interval −1.27% to 0.54%). No significant differences were found between the routes in the cumulative incidence of tuberculosis using a range of equivalence of "within 25%." Additionally, no significant differences were found between the routes in the cumulative incidence of adverse events (risk ratio 0.98, 95% confidence interval 0.91 to 1.06), including deaths (1.19, 0.89 to 1.58).

Conclusion Equivalence was found between intradermal BCG vaccine and percutaneous BCG in the incidence of tuberculosis in South African infants vaccinated at birth and followed up for two years. The World Health Organization should consider revising its policy of preferential intradermal vaccination to allow national immunisation programmes to choose percutaneous vaccination if that is more practical.

Trial registration ClinicalTrials.gov NCT00242047.

INTRODUCTION

The World Health Organization recommends giving BCG by the intradermal route.[1] Some countries prefer percutaneous BCG because of concerns about adverse events associated with the intradermal route.[2] As of 2000, Japan and South Africa were the only two countries with large BCG vaccination programmes still using percutaneous BCG.

In 1999, the South African National Department of Health decided to switch from percutaneous BCG to intradermal BCG in line with WHO recommendations. We compared the incidence of tuberculosis over two years in infants vaccinated at birth with intradermal BCG or with percutaneous BCG.

METHODS

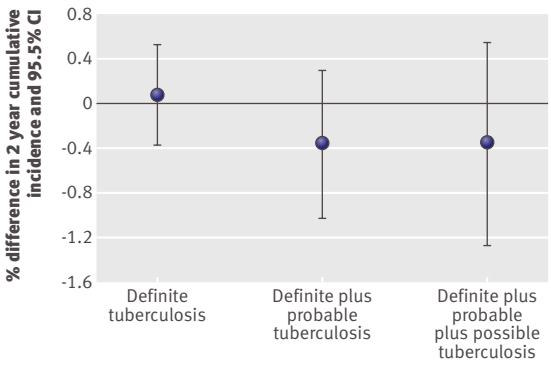
We carried out the study in a rural region of South Africa. In 2001 the area had an annual incidence of tuberculosis in adults of 515 per 100 000, and an estimated annual rate of tuberculosis in under 5s of 2.5%.[3 4]

We chose the Tokyo 172 BCG substrain vaccine (BCG Laboratory, Tokyo, Japan) because it is available in percutaneous and intradermal formulations. Intradermal BCG was given in a dose of 20×10^6 colony forming units/ml and percutaneous BCG as one drop of solution, containing 3×10^9 colony forming units/ml, through the skin in two places using a nine pronged tool (BCG Laboratory). The infants were vaccinated within 24 hours of birth.

Our primary outcome was the incidence of tuberculous events with bacteriological or histological confirmation or that met clinical criteria occurring within two years of birth. Secondary outcomes were the rates of adverse events, admissions to hospital for all causes and for tuberculosis, and mortality.

We admitted infants with suspected tuberculosis before age 2 years to a ward for verification. A study doctor and, when appropriate, study nurses carried out a clinical examination, chest radiography, tuberculin skin testing, HIV testing, and gastric aspiration and sputum induction for mycobacterial culture. Confirmed tuberculosis was categorised as definite, probable, or possible using a diagnostic algorithm (see bmj.com for definitions).

The null hypothesis was that there would be a difference in the incidence of tuberculosis between the two groups of at least Δ, where Δ equalled 25%. A range of equivalence of within 25% was prespecified based on



Percentage difference (95.5% confidence intervals) in cumulative incidence of tuberculosis over two years by end point

interpretation of a clinically meaningful difference associated with the routes. We calculated that a cohort of 12 000 infants, randomised in a 1:1 ratio and followed up for two years, would yield enough cases in each arm to assess equivalence in the range 2.5% (SD 25%) per annum with a power (1−β) of 0.80 and at a level of significance (α) of 0.05.

Infants were randomly allocated to route of BCG vaccine according to the week of birth. Participants were enrolled by midwives and study counsellors based at maternity units who were aware of the week's allocation.

Statistical analysis

The primary end point was the cumulative two year incidence of tuberculosis. The numerator was all infants meeting the criteria for tuberculosis; the denominator all randomised infants. Each infant was counted once only and in the case of multiple referrals for evaluation for tuberculosis we selected the most severe diagnosis. We calculated exact binomial confidence intervals for the cumulative incidence rates. To account for two partially unblinded interim analyses, which used 0.0005 and 0.014% significance levels, respectively, we used a 4.5% significance level in the final analysis, following a previously published method.[5] Equivalence was shown if the 95.5% confidence intervals for the incidence rates in the percutaneous group fell within 25% of those for the intradermal group. We further compared the cumulative incidence rates for the groups using risk ratios and corresponding Fisher's exact confidence interval.

We used cumulative incidence proportion and exact 95% binomial confidence interval to summarise the number of infants with at least one serious adverse event, the number of grade 3 and 4 vaccine related adverse events, and the number of deaths and compared these using risk ratios with corresponding 95% confidence intervals.

RESULTS

In total, 11 680 infants were randomised and vaccinated with BCG—5905 intradermally and 5775

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2052

percutaneously These infants were passively followed up for tuberculosis, morbidity, and mortality and were included in the primary intention to treat analyses. Two hundred and thirty protocol violations were recorded. These infants were included in the intention to treat analysis.

Infants were recruited from 26 March 2001 to 31 July 2004. The follow-up period lasted until 31 July 2006 or until the infant reached 2 years of age, whichever occurred first.

The vaccinated infants contributed person time to the denominators. Loss to follow-up was assumed to be evenly split between routes. It was not adjusted for in the calculation of incidence rates. All infants were analysed in the group to which they has been randomised, regardless of vaccine received. The primary outcome of the trial was documented tuberculosis or radiological and clinical evidence of the disease, expressed as the cumulative incidence of definite, probable, or possible tuberculosis over the two years.

Equivalence was found between the routes for definite, probable, or possible tuberculosis (figure). This is because the bounds of the 95.5% confidence intervals around the point estimate of the rate in the comparator group (percutaneous) lay within the interval bounded by the point estimate of the rate in the reference group (intradermal) plus or minus 25% (4.58% to 7.64%).[6]

Few cases of disseminated tuberculosis were detected and there were no statistically significant differences by route of vaccination. Three cases of tuberculous meningitis occurred in the intradermal group and one in the percutaneous group (risk difference 0.03%, 95% confidence interval −0.03% to 0.10%). No cases of miliary tuberculosis occurred in the intradermal group but one occurred in the percutaneous group (risk difference −0.02%, −0.05% to 0.02%). One case of abdominal tuberculosis and one case of spinal tuberculosis occurred in the percutaneous group.

Only those infants with suspected tuberculosis and admitted to the ward for investigations were tested for HIV. Thirty one of 1576 (1.97%) tested positive; 9.7% had definite tuberculosis, 35.5% probable tuberculosis, and 16.1% possible tuberculosis. Overall, 2180 infants were admitted to hospital: cumulative incidence 18.7% (1091 (18.5%) in intradermal group, 1089 (18.9%) in percutaneous group), risk ratio 0.98 (95% confidence interval 0.91 to 1.06).

Twenty one grade 3 and one grade 4 adverse reactions occurred (see bmj.com). One infant in the percutaneous group was diagnosed as having disseminated disease (subtype *M bovis* BCG substrain Tokyo 172). The infant was not infected with HIV.

One hundred and two (1.73%) infants in the intradermal group and 84 (1.45%) in the percutaneous group died: risk ratio (1.19, 0.89 to 1.58): cumulative incidence 1.59%. None of these deaths was considered to be related to the intervention.

WHAT IS ALREADY KNOWN ON THE TOPIC

Vaccination of newborn infants with BCG provides variable and incomplete protection against tuberculosis

WHO and most countries advocate the intradermal route for administering BCG vaccine

No study to date has provided clear evidence that one route of vaccination is safer, more immunogenic, or more efficacious than the other

WHAT THIS STUDY ADDS

The intradermal route for administering BCG vaccine to infants and young children was as efficacious as the percutaneous route for preventing tuberculosis

The two routes were equivalent in terms of safety

DISCUSSION

In a randomised trial we found equivalence in the cumulative incidence of definite, probable, and possible tuberculosis in infants before age 2 years who received BCG through either the intradermal route or the percutaneous route at birth.

The diagnosis of tuberculosis in infants is difficult. The most robust definition of tuberculosis in this age group is definite or probable. This was not used as the primary outcome in this trial. Tuberculosis in young children in countries with a high burden of tuberculosis is generally not diagnosed from cultures owing to unavailability.

For a variety of reasons intradermal BCG is widely used and is the method advocated by WHO.[1] Although evidence is scant, a 2004 review of studies comparing the intradermal route with the percutaneous route found that percutaneous BCG is less efficient in stimulating interferon-γ production and suggested that intradermal BCG should be recommended.[1] This route is, however, technically challenging and requires intensive quality control.

The rates of tuberculosis in our study were high. Neither vaccination route worked particularly well in preventing pulmonary disease, with over 6% of infants developing definite, probable, or possible tuberculosis. Even if possible cases were discounted, the rate is over 3%.

We found few cases of disseminated disease. We detected four cases of tuberculous meningitis, giving a cumulative incidence of 0.03% or 0.015% per year, half of that expected. Miliary tuberculosis is said to occur in about 2.7% of children with active tuberculosis.[7] Given our calculated cumulative incidence of all tuberculosis over two years of 6.31%, the cumulative incidence of miliary tuberculosis should have been 0.17% over two years, or 20 cases, but we saw only one case, giving a cumulative incidence of 0.009%, about 20 times less than expected.

There are several possible explanations for this. Firstly, the process of doing the trial brought resources into the area that had not been available previously.

Some of these were used to trace contacts of adults with tuberculosis, evaluate them for tuberculosis, and refer them for prophylaxis or treatment . Many infants were therefore detected soon after they had been infected, before they could develop clinical disease. Secondly, we might have missed cases and deaths due to disseminated disease because of suboptimal detection. We think it is unlikely, however, as health providers in South Africa are well aware of the condition. Thirdly, we identified infants with subclinical disease but positive culture results, some of whom might have self cured. Lastly, our sample was biased towards more healthy infants.

Overall, 186 infants died. A mortality surveillance system was developed to identify the causes as far as possible. Pneumonia, gastroenteritis, and septicaemia were among the commonest causes, and important underlying causes included HIV/AIDS, prematurity, low birth weight, and malnutrition. Mortality was largely due to infectious diseases. Many deaths could only be categorised as "sudden unexplained" or as "ill defined," illustrating the difficulty of accurate surveillance in field trials in developing countries. The number of adverse vaccine reactions was relatively low.

We were not able to show equivalence in the cumulative incidence of definite and probable tuberculosis or of definite tuberculosis only between the groups, and in the definite only group the difference was in the opposite direction, with marginally lower rates in the percutaneous group. The study was powered to detect a difference using the most sensitive and least specific end point but had our numbers been larger we would have shown equivalence for both the alternative, more specific end points.

Study limitations

Eleven of 11 691 women gave consent before but not after delivery. Also, it was not possible to blind staff assessing outcomes; chest radiography and microbiological tests were carried out and reviewed blinded, as were the final categorisation of cases. It is possible that we overestimated the absolute rates of culture positive tuberculosis as we cannot exclude cross contamination. The caregivers' recall of exposure to contacts may be inaccurate. The significance of a tuberculin skin test result is difficult to gauge in a population with malnutrition and other causes of suppressed cellular immunity, BCG vaccination is close to universal, and exposure to non-tuberculous mycobacteria common. The clinical diagnosis of paediatric tuberculosis is subjective as the symptoms and signs in children are non-specific. We enrolled about two thirds of the region's birth cohort. Those not enrolled might have come from poorer families and thus have been more likely to have tuberculosis. We included only infants who were eligible for routine BCG vaccination at birth. Our results may therefore be less generalisable, specifically to groups at higher risk of tuberculosis. We compared two routes for giving BCG substrain Tokyo 172. It is possible that some strains are more efficacious than others in preventing tuberculosis.[2 8-10]

One hundred and sixteen (63 intradermal group, 53 percutaneous group) children were diagnosed as having tuberculosis outside of the case verification ward. Children who had a diagnosis of tuberculosis during a routine hospital admission were likely to have more advanced disease, and more missing diagnostic data. We therefore included in the primary analysis only those children with tuberculosis diagnosed on the ward. Despite all the potential limitations of the study, we believe that none should have significantly influenced the derived relative efficacy estimates.

Data Safety Monitoring Committee: Prakash Jeena (chair), Christine Grady, Lesley Henley, Linda Gail Bekker, Lisa Saiman, Mike Stoto, Neil Cameron, and Nicol Coetzee. Expert clinical panel: Maurice Kibel, John Burgess, and Peter Donald. Expert radiology panel: John Burgess, Maurice Kibel, and Robert Gie. The South African BCG trial team consisted of Andre Burger, Angela Malone, Anne Marie Demers, Ashley Veldsman, Aysel Gumusboga, Carolyn Kewley, Deon Minnies, Desire Michaels, Elmarie Simon, Gilla Kaplan, Heather Zar, Karen Iloni, Lea Denation, Lesley Workman, Linda van der Merwe, Marcel Behr, Marie Buchanan, Martien Borgdorff, Marijke Geldenhuys, Maurice Kibel, Michele Tameris, Monique Hanslo, Ronel Shepherd, Simon Schaaf, Siziwe Mawu, Sizulu Moyo, Stefanie Abraham, Suzanne Verver, Sylvia Mlanjeni, and Veronica Dirks. We thank Takeru Hashimoto, Akira Koyama, Naoki Nakada, and the management of the Tokyo BCG laboratory, Japan; Frans Krige and management of health department, Boland Overberg Region, Lizette Phillips and management of Brewelskloof Hospital, Worcester, and the participants.

Contributors: See bmj.com.

Funding: The study was supported by the Aeras Global TB Vaccine Foundation, Rockville, MD, USA, a non-profit organisation that aims to develop tuberculosis vaccines. LG (co-principal investigator) and LB, JS, and MAG (co-investigators) were fulltime employees of the sponsor at the time of the trial.

Competing interests: None declared.

Ethical approval: The study protocol was approved by the research ethics committee of the Faculty of Health Sciences, University of Cape Town and the research in human subjects committee of the Aeras Global TB Vaccine Foundation. The trial was not externally monitored but two audits were done during the enrolment period, on behalf of the sponsor, by an independent contract research organisation, Triclinium, Johannesburg, South Africa.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Bricks LF. Percutaneous or intradermal BCG vaccine? *J Pediatr (Rio J)* 2004;80:93-8.
2 Al Jarad NA, Empey DW, Duckworth G. Administration of the BCG vaccination using the multipuncture method in school children: a comparison with intradermal method. *Thorax* 1999;54:762-4.
3 Municipal Demarcation Board, South Africa. 2007. www.demarcation.org.za.
4 Groenewald P. Boland Overberg Region annual report 2002. Worcester, South Africa: Department of Information Management, Boland Overberg Regional Office.
5 O'Brien P, Fleming T. A multiple testing procedure for clinical trials. *Biometrics* 1979;35:549-56.
6 Pater C. Equivalence and noninferiority trials—are they viable alternatives for registration of new drugs? *(III) Curr Control Trials Cardiovasc Med* 2004;5:8.
7 Starke JR, Taylor-Watts KT. Tuberculosis in the pediatric population of Houston, Texas. *Pediatrics* 1989;84:28-35.
8 Ormerod LP, Palmer C. Tuberculin reactivity after neonatal percutaneous BCG immunization. *Arch Dis Child* 1993;69:155.
9 Hart PD. Efficacy and applicability of mass BCG vaccination in tuberculosis control. *BMJ* 1967;1:587-92.
10 Kemp EB, Belshe RB, Hoft DF. Immune responses stimulated by percutaneous and intradermal bacille Calmette-Guérin. *J Infect Dis* 1996;174:113-9.

Accepted: 4 September 2008

Inability to get up after falling, subsequent time on floor, and summoning help: prospective cohort study in people over 90

Jane Fleming, Carol Brayne, and the Cambridge City over-75s Cohort (CC75C) study collaboration

EDITORIAL by Campbell

Department of Public Health and Primary Care, University of Cambridge, Institute of Public Health, Cambridge CB2 0SR

Correspondence to:
jane.fleming@phpc.cam.ac.uk
carol.brayne@medschl.cam.ac.uk

Cite this as: *BMJ* 2008;337:a2227
doi:10.1136/bmj.a2227

ABSTRACT

Objectives To describe the incidence and extent of lying on the floor for a long time after being unable to get up from a fall among people aged over 90; to explore their use of call alarm systems in these circumstances.

Design 1 year follow-up of participants in a prospective cohort study of ageing, using fall calendars, phone calls, and visits.

Setting Participants' usual place of residence (own homes or care homes), mostly in Cambridge.

Participants 90 women and 20 men aged over 90 (n=110), surviving participants of the Cambridge City over-75s Cohort, a population based sample.

Main outcome measures Inability to get up without help, lying on floor for a long time after falling, associated factors; availability and use of call alarm systems; participants' views on using call alarms to summon help if needed after falling.

Results In one year's intensive follow-up, 54% (144/265) of fall reports described the participant as being found on the floor and 82% (217/265) of falls occurred when the person was alone. Of the 60% who fell, 80% (53/66) were unable to get up after at least one fall and 30% (20/66) had lain on the floor for an hour or more. Difficulty in getting up was consistently associated with age, reported mobility, and severe cognitive impairment. Cognition was the only characteristic that predicted lying on the floor for a long time. Lying on the floor for a long time was strongly associated with serious injuries, admission to hospital, and subsequent moves into long term care. Call alarms were widely available but were not used in most cases of falls that led to lying on the floor for a long time. Comments from older people and carers showed the complexity of issues around the use of call alarms, including perceptions of irrelevance, concerns about independence, and practical difficulties.

Conclusions Lying on the floor for a long time after falling is more common among the "oldest old" than previously thought and is associated with serious consequences. Factors indicating higher risk and comments from participants suggest practical implications. People need training in strategies to get up from the floor. Work is needed on access and activation issues for design of call alarms and information for their effective use. Care providers need better understanding of the perceptions of older people to provide acceptable support services.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2227

INTRODUCTION

The risk of falling increases with age and ability to get up declines, so the risks of any fall are far greater because of the complications that can ensue from lying on the floor for a long time—for example, pressure sores, dehydration, hypothermia, pneumonia, and even death.[1] This inability to get up has a poor prognosis in terms of the risk of injury in a subsequent fall,[2] admission to hospital,[1] and mortality.[1 3]

Personal emergency response systems might be cost effective in reducing hospital admissions,[4 5] but uptake and adherence are low.[6 7] In our prospective study of falls among the oldest old (people aged ≥90) we sought to quantify the numbers lying on the floor a long time and explored the extent to which alarms were used.

METHODS

Data were collected on the immediate consequences of falls among participants, all aged over 90, of a population based study, the Cambridge City over-75s Cohort (CC75C). Full methods are described elsewhere (www.cc75c.group.cam.ac.uk).[8-10] Participants were followed up for one year or until death if sooner. Falls were reported by fall calendars posted weekly or by telephone with follow-up visit or phone call by the project nurse. Information was sought both from the participant and any proxy informant. We used the term "severe cognitive impairment" for those with scores 0-17 on the mini-mental state examination[11] or diagnosed dementia.

RESULTS

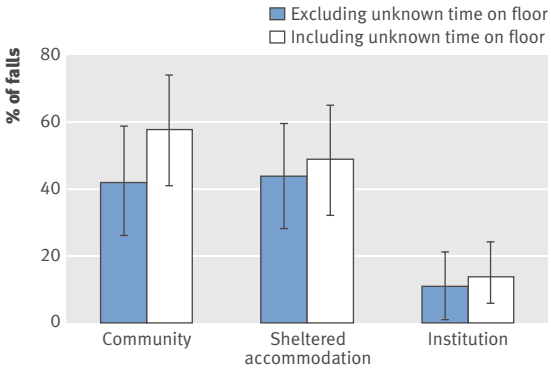
Sample characteristics

Of the surviving CC75C study participants, 110 (84%), 90 women and 20 men, took part in the falls survey and follow-up (age 91-105, median 94). Sixty six (60%) were reported to have fallen at least once during the year after interview.[9 10] Most were unable to climb stairs, a third were unable to walk outdoors, about a third were severely cognitively impaired, and one in 10 was housebound. Two thirds of the people living in the community lived alone, fewer than half were living in institutional or sheltered settings, and more than a fifth were still able to walk around their local neighbourhood.

Falling

During follow-up at least one fall was reported for 56% (35/62) of the participants living in the community, 68%

(13/19) of those living in sheltered housing, and 62% (18/29) of those in institutional settings (residential care, nursing home, or hospital) at the time of interview. There were 265 reported falls in total as most people fell at least twice.[10] Falls in the community accounted for 45% (120/265) of all falls reported; the remaining 23% (62/265) and 31% (83/265) of falls happened in sheltered housing schemes and institutional settings respectively. Over half (54%, 144/265) of fall reports described the participant as being found on the floor. The proportion of falls that happened when the person who fell was on their own was 82% (217/265) overall but as high as 94% (58/62) in sheltered accommodation. Of the people who fell, rather than falls, 88% (58/66) were alone when at least one of these falls happened.



Lying on the floor for an hour or more after falling (143 falls by people aged >90 who were alone and unable to get up). Proportions with 95% confidence intervals

Inability to get up

Everyone who fell in an institutional setting, 66% (41/62) who fell in sheltered housing, and 43% (52/120) who fell in community settings needed help to get up after a fall. In two thirds of all the falls reported the person who fell was unable to get up without help (176/265, 66%), but 80% (53/66) of the participants who fell had difficulty getting up from at least one fall.

Fifteen per cent (n=40) of all reported falls resulted in the person lying on the floor for an hour or more. When we considered people who fell, rather than falls, these proportions were even higher: 30% (n=20) of those who fell were on the floor for an hour or more on at least one occasion. For a further 6% (n=16) of falls and 9% of people who fell (n=6) the maximum time was unknown.

The length of time spent on the floor after falling depended both on ability to get up and on whether there was any help at hand. Overall 28% (40/143) of unwitnessed falls led to the person lying on the floor for an hour or more. This is lower than the high prevalence in such situations among residents in sheltered housing (44%) and participants living in the community (42%) (figure).

Factors associated with inability to get up and lying on the floor for a long time

The table describes the participants who were unable to get up after a fall (n=53) and those who were on the floor for at least an hour (n=20) on at least one occasion during the follow-up year.

Women were six times more likely than men to have difficulty getting up. Residents of any supported living setting had a 16-fold increased odds of being unable to get off the floor without help; no one who fell in a care home was able to get up unaided. Those who lived alone in the community or in sheltered housing were particularly unlikely to be able to get up after falling. Reported mobility was consistently associated with inability to get up, and climbing less than a flight of stairs a day was the most consistent predictor for this serious consequence of falls, outweighing the effects of cognitive impairment and recalled falls.

Severe cognitive impairment was significantly associated with lying on the floor for an hour or more. Living alone quadrupled the odds of lying on the floor for a long time. No other descriptive characteristics predicted the length of time on the floor. Of the 20

Attitudes towards the use of call alarms to summon help

Not having a call alarm

"My niece is only next door. I can bang on the wall if I need to call help."

Daughter: "She refuses to have a call alarm because she thinks it would keep going off by mistake. She is worried enough about the string pull alarms in each room [sheltered housing scheme] and often won't turn on the kitchen or bathroom lights in case she pulls the wrong cord."

Having one but not wearing it

"I don't have to wear it yet, I just hang it on the back of the chair"

"I wasn't wearing my pendant. I don't usually wear it . . . It was quite a struggle to get up. It took about half an hour."

"I'd already taken it off ready for bed and put it on the bedside table so then I couldn't reach it." [Lost balance undressing, on the floor until next morning]

Wearing but choosing not to use it

"I wanted to be able to get up by myself. It took me a long time to get up but I did it in the end. It makes me annoyed if I have to have help."

"I didn't want to use the call alarm, although I was wearing it, for fear of being taken into hospital."

"Didn't need to." [despite falling twice]

Difficulty in activating alarm

"I couldn't have reached the alarm."

"I tried to call Care Call but the pendant didn't work because there'd been a power cut the day before."

"I was wearing my alarm but I didn't think it would work out there in the street so I didn't even try it."

"It always seems a long time when you're waiting but I don't suppose it was really." [Found on floor by carer, confused. She had not set off the alarm, although she thought she had]

people who were reported to be on the floor over an hour after falling, 60% (12) had a fall related hospital admission during the follow-up year (adjusted odds ratio 4.0, 95% confidence interval 1.3 to 12.3). Thirty nine per cent (15/38) of those who had been unable to get up from the floor and 53% (8/15) of those who were on the floor over an hour had moved into long term care by the end of the study.

Use of call alarm systems to summon help

Many of those in the study population had call alarm systems. Of those living alone in the community or sheltered accommodation, 70% (57/81) had some form of call alarm, either personally worn (47%), or in the room (7%), or both (16%).

In 99% (141/143) of falls in those who could not get up when they fell alone the person had some form of call alarm system. In 80% (113/141) of these falls, however, they did not use their call alarm to summon help. Of these 141 falls, 38 resulted in lying on the floor for an hour or more, despite an installed alarm system, and in 97% of these "long lies" (37/38) the person who fell alone did not use their alarm to summon help. Lying on the floor for an hour or more occurred in only one of the 28 falls in which a call alarm was activated.

Factors influencing use of call alarms

Several themes emerged from comments made by participants and carers about why call alarms were not used. The box illustrates the range of attitudes.

DISCUSSION

Study summary

We found high rates of serious consequences of falling —being unable to get up and thus lying on the floor for a long time, the latter strongly associated with cognitive impairment. Lying on the floor for an hour or more was also strongly associated with serious injuries, admissions to hospital, and subsequent moves into long term care. Call alarm systems were widely available but were often not used.

Limitations

Even in this prospective data collection we had to rely on recall to some extent. We maximised validity of the

Factors associated with inability to get up and lying on the floor for a long time after fall, with unadjusted and adjusted odds ratios (95% confidence intervals)

	Unable to get up without help after ≥1 fall		Lying on floor for at least 1 hour after ≥1 fall	
	No of participants (n=53)	Adjusted* OR (95% CI)	No of participants (n=20)	Adjusted* OR (95% CI)
Mean (SD) age (years)	94.9 (2.8)		94.5 (1.9)	
For each additional year of age	—	1.7 (1.0 to 2.8)	—	1.0 (0.8 to 1.2)
Women *v* men	48 (91)	2.5 (0.4 to 15.6)	18 (90)	1.5 (0.3 to 8.7)
Place of residence:				
Supported setting (sheltered/institution) *v* community	30 (57)	7.1 (0.6 to 76.1)	10 (50)	0.9 (0.3 to 3.0)
Living alone *v* with spouse/other family (excluding institutions)	31/35 (89)	10.1 (1.0 to 99.0)	13/14 (93)	5.9 (0.5 to 77.6)
Maximum walking distance:				
Unable *v* able to walk around local area	46 (87)	5.0 (1.0 to 26.7)	18 (90)	3.0 (0.6 to 15.9)
Use of walking aid:				
Needed *v* not needed outdoors (excluding cannot walk outdoors)	44/46 (96)	2.4 (0.2 to 23.6)	17/18 (94)	1.4 (0.1 to 15.5)
Needed *v* not needed indoors	39 (74)	1.7 (0.3 to 8.7)	16 (80)	2.6 (0.6 to 10.3)
Cannot climb stairs or <1 flight/day *v* ≥1 flight/day	44 (83)	16.6 (3.1 to 87.7)	16 (80)	1.7 (0.5 to 6.7)
Severe cognitive impairment† *v* moderate/no impairment	23 (43)	4.1 (0.4 to 46.3)	13 (65)	8.1 (2.1 to 31.0)
Recalled falls in past year‡:				
≥1 *v* none	37 (71)	2.7 (0.5 to 13.9)	13 (65)	0.9 (0.3 to 2.9)
Injuries sustained in falling:				
Any *v* none	40 (75)	3.0 (0.6 to 14.6)	17 (85)	3.5 (0.9 to 14.4)
Serious injury v non-serious/no injury	13 (25)	3.2 (0.3 to 42.7)	8 (40)	4.2 (1.2 to 14.8)
Serious *v* none (excluding non-serious)	13/26 (50)	6.6 (03. to 170.8)	8/11 (73)	7.4 (1.3 to 41.1)
Hospital admissions:				
Fall related hospital admission *v* no fall related admission	24 (45)	21.1(1.9 to 230.5)	12 (60)	4.0 (1.3 to 12.3)

*Model selected by multiple variable regression adjusted for age, sex, and mobility (climbing less than flight of stairs/day); hospital admissions adjusted only for mobility because of low numbers).
†Mini-mental state examination score 0-17 or diagnosis of dementia.
‡Data missing for 1.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Being unable to get up after falling can have serious consequences for an old person, especially if they are on the floor a long time

WHAT THIS STUDY ADDS

Being unable to get up and therefore lying on the floor for a long time are prevalent after falls in men and women over 90

Severe cognitive impairment was the only intrinsic factor predicting lying on the floor for a long time

Lying on the floor for a long time after a fall was associated with repeated falls, fall induced injuries, and subsequent admissions to hospital and long term care

Alarm systems were widely available but rarely used to call for help to get up

data by using a combination of methods—participant and proxy calendar reports, phone calls, and visiting. There was little scope for misreporting the proportion of people unable to get up from the floor unaided, as they were on the floor when help arrived, but length of time on the floor is a less robust measure. Caution is warranted in interpreting the analyses of association between risk factors and these consequences of falling; lack of association could be due to the limited power of the small sample size.

Implications

Reported limitations with mobility—particularly on stairs—were strongly associated with being unable to get up after falling, suggesting that those at risk could be readily identified for preventative initiatives such as training in how to get up, for which there is growing evidence.[12] Cognitive impairment was the only characteristic that predicted lying on the floor for a long time, probably indicating that the most cognitively impaired are the least likely to summon help when they cannot get up. The development of automatic fall detectors that do not rely on the wearer activating them[13] could help to reduce time on the floor. Difficulty in getting up from the floor is more likely in people in supported living but the odds of lying on the floor for a long time were not reduced by the additional level of care. Calling for help can still be problematic in institutional settings.

Nearly all people who couldn't get up had access to a call alarm system but did not use it to summon help, raising important questions for care providers. Few studies to date have explored older people's views on assistive technology devices,[6 7] some reporting overall positive attitudes,[14] but older people are understandably reluctant to be labelled as at high risk for falling.[15] The emergent themes in this study include older people's justifiable concern to preserve their independence. There are practical implications arising from the comments from some of the frailer individuals who tried unsuccessfully to use their alarms.

Comments from study participants and their relatives revealed pertinent concerns on which further qualitative research might shed more light, particularly to inform the development of interventions to reduce the length of time people lie on the floor after a fall, including the design of call or detector systems. These are complex issues that care providers need to understand better to offer and effectively deliver support services and devices that are acceptable to the older people concerned.

We gratefully acknowledge the contributions of previous investigators and past research team members (see www.cc75c.group.cam.ac.uk/pages/studypersonnel), collaborating general practitioners and their practice teams, and, most particularly, the study respondents, their families, and friends and the staff in many care homes. Rosemary Abbott, Stephen Barclay, Tom Dening, Morag Farquhar, Felicia Huppert, Ann-Louise Kinmonth, Fiona Matthews, Elizabeta Mukaetova-Ladinska, Graciela Muniz, Angela O'Sullivan, Eugene Paykel, Kathryn Richardson, and Emily Zhao (current CC75C collaborators) commented on earlier drafts of this paper.

Contributors: See bmj.com.

Funding: This study was supported by an NHS executive research and development unit health services research fellowship and CC75C is currently supported by the BUPA Foundation. CC75C is a study within the Cambridgeshire and Peterborough Collaboration for Leadership in Applied Health Research and Care. No funders (see http://www.cc75c.group.cam.ac.uk/pages/grant for full list) had any role in the study design, analysis, or writing of this paper.

Competing interests: None declared.

Ethical approval: For each participant written informed consent was obtained either in person or from the proxy informant, as approved by Cambridge research ethics committee.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Tinetti ME, Liu WL, Claus EB. Predictors and prognosis of inability to get up after falls among elderly persons. *JAMA* 1993;269:65-70.
2 Bergland A, Laake K. Concurrent and predictive validity of "getting up from lying on the floor". *Aging Clin Exp Res* 2005;17:181-5.
3 Wild D, Nayak US, Isaacs B. How dangerous are falls in old people at home? *BMJ* 1981;282:266-8.
4 Roush RE, Teasdale TA, Murphy JN, Kirk MS. Impact of a personal emergency response system on hospital utilization by community-residing elders. *South Med J* 1995;88:917-22.
5 Bernstein M. "Low-tech" personal emergency response systems reduce costs and improve outcomes. *Manag Care Q* 2000;8:38-43.
6 Levine DA, Tideiksaar R. Personal emergency response systems: factors associated with use among older persons. *Mt Sinai J Med* 1995;62:293-7.
7 Mann WC, Belchior P, Tomita MR, Kemp BJ. Use of personal emergency response systems by older individuals with disabilities. *Assist Technol* 2005;17:82-8.
8 Fleming J, Zhao E, O'Connor DW, Pollitt PA, Brayne C, CC75C study collaboration. Cohort profile: the Cambridge city over-75s cohort (CC75C). *Int J Epidemiol* 2007;36:40-6.
9 Fleming J. Patterns and predictors of falls and their consequences in extreme old age. [PhD thesis.] Cambridge: University of Cambridge, 2006.
10 Fleming J, Brayne C, CC75C collaborative. Falls in advanced old age: recalled falls and prospective follow-up of over-90-year-olds in the CC75C study. *BMC Geriatrics* 2008;8:6..
11 Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189-98.
12 Skelton D, Dinan S, Campbell M, Rutherford O. Tailored group exercise (falls management exercise—FaME) reduces falls in community-dwelling older frequent fallers (an RCT). *Age Ageing* 2005;34:636-9.
13 Lindemann U, Hock A, Stuber M, Keck W, Becker C. Evaluation of a fall detector based on accelerometers: a pilot study. *Med Biol Eng Comput* 2005;43:548-51.
14 Demiris G, Rantz M, Aud M, Marek K, Tyrer H, Skubic M, et al. Older adults' attitudes towards and perceptions of "smart home" technologies: a pilot study. *Med Inform Internet Med* 2004;29:87-94.
15 Yardley L, Bishop FL, Beyer N, Hauer K, Kempen GI, Piot-Ziegler C, et al. Older people's views of falls-prevention interventions in six European countries. *Gerontologist* 2006;46:650-60.

Accepted: 5 September 2008

Recruitment to multicentre trials—lessons from UKCTOCS: descriptive study

Usha Menon,[1] Aleksandra Gentry-Maharaj,[1] Andy Ryan,[1] Aarti Sharma,[1] Matthew Burnell,[1] Rachel Hallett,[1] Sara Lewis,[1] Alberto Lopez,[2] Keith Godfrey,[2] David Oram,[3] Jonathan Herod,[4] Karin Williamson,[5] Mourad Seif,[6] Ian Scott,[7] Tim Mould,[8] Robert Woolas,[9] John Murdoch,[10] Stephen Dobbs,[11] Nazar Amso,[12] Simon Leeson,[13] Derek Cruickshank,[14] Ali McGuire,[15] Stuart Campbell,[16] Lesley Fallowfield,[17] Steve Skates,[18] Mahesh Parmar,[19] Ian Jacobs[1]

[1]Gynaecological Oncology, UCL EGA Institute for Women's Health, London W1T 7DN
[2]Queen Elizabeth Hospital, Gateshead NE9 6SX
[3]St Bartholomew's Hospital, London EC1A 7BE
[4]Liverpool Women's Hospital, Liverpool L8 7SS
[5]Nottingham City Hospital, Nottingham NG5 1PB
[6]St Mary's Hospital, Manchester M13 9WL
[7]Derby City Hospital, Derby DE22 3NE
[8]Royal Free Hospital, London NW3 2QG
[9]St Mary's Hospital, Portsmouth PO3 6AD
[10]St Michael's Hospital, Bristol BS2 8EG
[11]Belfast City Hospital, Belfast BT9 7AB
[12]University of Wales College of Medicine, Cardiff CF14 4XN
[13]Llandudno Hospital, Llandudno LL30 1LB
[14]James Cook University Hospital, Middlesbrough TS4 3BW
[15]London School of Economics, London WC2A 2AE
[16]Create Health Clinic, London W1G 6AJ
[17]Brighton and Sussex Medical School, University of Sussex, Brighton BN1 9PX
[18]Harvard Medical School, Boston, MA 02115, USA
[19]Medical Research Council Clinical Trials Unit, London NW1 2DA

Correspondence to: U Menon
u.menon@ucl.ac.uk

Cite this as: *BMJ* 2008;337:a2079
doi:10.1136/bmj.a2079

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2079

ABSTRACT

Objective To describe the factors that contributed to successful recruitment of more than 200 000 women to the UK Collaborative Trial of Ovarian Cancer Screening, one of the largest ever randomised controlled trials.

Design Descriptive study.

Setting 13 NHS trusts in England, Wales, and Northern Ireland.

Participants Postmenopausal women aged 50-74; exclusion criteria included ovarian malignancy, bilateral oophorectomy, increased risk of familial ovarian cancer, active non-ovarian malignancy, and participation in other ovarian cancer screening trials.

Main outcome measures Achievement of target recruitment, acceptance rates of invitation, and recruitment rates.

Results The trial was set up in 13 centres with 27 adjoining local health authorities. The coordinating centre team was led by one of the senior investigators, who was closely involved in planning and day to day trial management. Of 1 243 282 women invited, 23.2% (288 955) replied that they were eligible and would like to participate. Of those sent appointments, 73.6% (205 090) attended for recruitment. The acceptance rate varied from 19% to 33% between trial centres. Measures to ensure target recruitment included named coordinating centre staff supporting and monitoring each centre, prompt identification and resolution of logistic problems, varying the volume of invitations by centre, using local non-attendance rates to determine the size of recruitment clinics, and organising large ad hoc clinics supported by coordinating centre staff. The trial randomised 202 638 women in 4.3 years.

Conclusions Planning and trial management are as important as trial design and require equal attention from senior investigators. Successful recruitment needs constant monitoring by a committed proactive management team that is willing to explore individual solutions for different centres and use central resources to improve local recruitment. Automation of trial processes with web based trial management systems is crucial in large multicentre randomised controlled trials. Recruitment can be further enhanced by using information videos and group discussions.

Trial registration Current Controlled Trials ISRCTN22488978.

INTRODUCTION

Only 31% of multicentre randomised controlled trials funded by the UK Medical Research Council and the NHS Health Technology Assessment Programme that were recruiting between 1994 and 2002 achieved their original recruitment target.[1] In 2000 expertise accumulated over a decade of planning and running ovarian cancer screening trials was used to design and set up the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS; www.ukctocs.org.uk).[2-6] Recruitment of the required target of 202 638 women was completed in 2005, making it one of the largest ever randomised controlled trials. This report describes the approach to planning and management that contributed to successful recruitment.

METHODS

The UKCTOCS involves 200 000 women randomised to annual screening with serum CA 125 or transvaginal ultrasound or no intervention. A senior investigator leads trial management based at the coordinating centre. We use a custom built, web based trial management system to centralise and automate trial processes. We identified NHS trusts (trial centres) wishing to participate and set them up in a staggered fashion over the course of two years. Recruitment started at a trial centre when at least 1500 local women had accepted the invitation. The launch of the trial was accompanied by national media coverage. This was followed by local media coverage in the form of radio interviews and newspaper articles as each centre started recruitment. We briefed the staff manning the telephone lines of the patient support charities OVACOME and Cancer BACUP and provided them with answers to frequently asked questions before any publicity.

Invitation

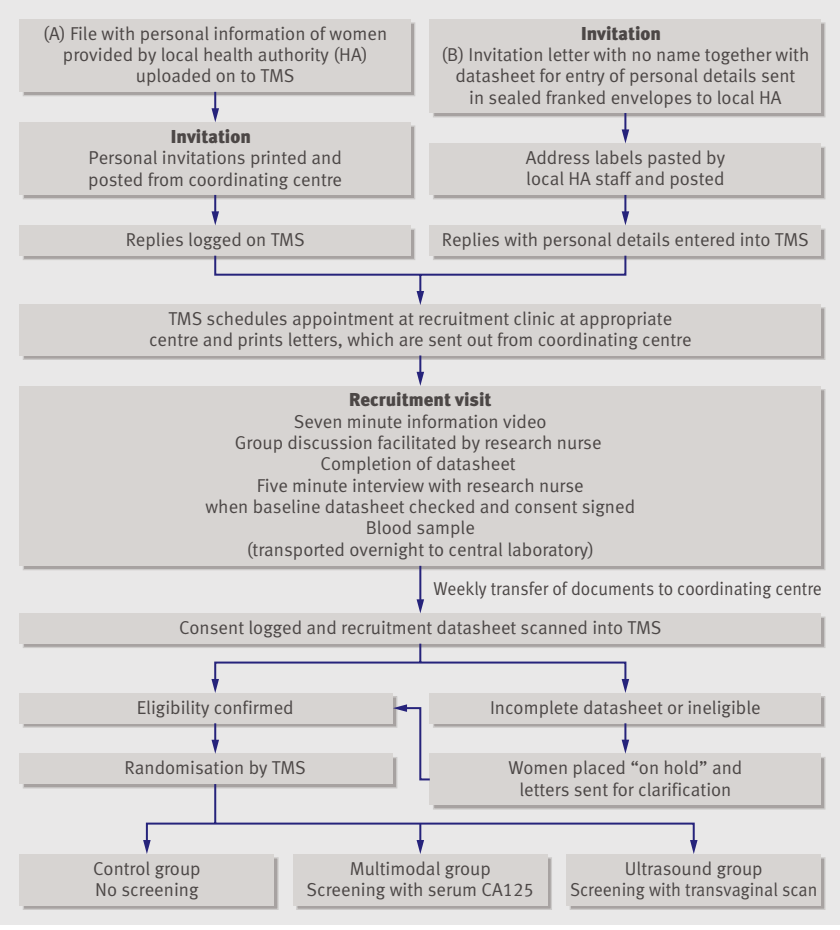
We sent information about the trial to all general practitioners working in participating primary care trusts, after electronic upload of their details into the trial management system. We requested electronic files containing details of 2000 to 10 000 women on a regular (usually three monthly) basis from each of the participating primary care trusts for upload. We then

sent women personal invitations and logged replies on the trial management system (figure). The patient support groups OVACOME and Cancer BACUP vetted the information for patients, and invitation letters contained their contact details. In the course of the trial, we revised and simplified the invitation.

Recruitment and randomisation

We set the weekly recruitment target at 100 women per trial centre. We set up individual profiles comprising five recruitment clinics a week on the trial management system for each trial centre. These consisted of 45 minute appointments involving groups of four women. Once acceptance was logged, women were automatically scheduled into the appointment slots at the centre and sent letters. The web browser enabled immediate logging of clinic attendance as well as rescheduling of appointments by both local and coordinating centre staff. At recruitment, women viewed an information video and participated in a group discussion. This was immediately followed by a one to one discussion in a separate room with the research nurse, when women were given the opportunity to discuss any private concerns before signing consent. The nurse also checked their completed datasheet. These documents were sent weekly to the coordinating centre, where the trial management system automatically confirmed eligibility, did randomisation, scheduled appropriate appointments to women allocated to screening, and printed letters to the patient and her general practitioner (figure). We sent requests for more information to women with incomplete data and placed "on hold" those whose last menstrual period was less than 12 months from date of recruitment. The trial management system also classifies screening results (ultrasound findings and CA 125 concentrations) entered over the web browser, schedules appropriate follow-up appointments, and prints all letters to individual women. It allows electronic exchange of information with the CA 125 analyser in the laboratory and with the Office for National Statistics.

We logged and monitored all complaints centrally. A designated person at the coordinating centre wrote to each woman directly after investigating the problems raised. We explored all suggestions by the women and amended trial logistics when appropriate and possible.



Invitation, recruitment, and randomisation. (A)=recruitment in which primary care trusts (PCTs) were allowed access to contact details of women; (B)=recruitment in which PCTs did not allow access to contact details of women; TMS=trial management system

RESULTS

Between April 2001 and February 2003 we set up the trial in 13 NHS trusts (trial centres) in England, Wales, and Northern Ireland (see bmj.com). Overall, 27 primary care trusts (including local health boards in Wales) with 3266 general practices were involved. Of these, 22 (81.5%) primary care trusts provided contact details of the women, and we sent invitations as outlined above. Five (18.5%) primary care trusts refused access to the contact details. For these trusts, we negotiated an alternative method using "unlinked" invitations. This involved sending the standard invitation letter without a recipient's name to the primary care trust in sealed franked envelopes. Trust staff then attached address labels and forwarded the letters. Women who wished to participate wrote to the coordinating centre with their contact details (figure).

Invitation—We closely monitored acceptance rates from the start and therefore noted within the first six months that recruitment was below target. We rapidly introduced measures to improve rates of acceptance of the invitation (box). Overall, invitations were sent to 1 243 282 women; 288 955 (23.2%) women replied that they would like to participate in the trial and were eligible. The overall acceptance rate varied between centres from 19% in East London to 33% in Bristol.

Recruitment and randomisation—Between April 2001 and September 2005 205 090 women (73.6% of those who were sent appointments) attended for recruitment, and we finally randomised 202 638 women). Ninety eight (0.008%) invited women complained about recruitment related problems: invitation to trial (32), trial information (28), recruitment appointment (17), and randomisation to control group (21). The number of women randomised each month ranged from 117 to

5773 (median 3955)[7]; the median time from recruitment to randomisation was 12.3 days (25th to 75th centile 8.5 to 15.5 days).

DISCUSSION

Our experience highlights the importance of meticulous planning and management of trial processes. We quickly learnt that the key to successful recruitment was constant monitoring by a dedicated management team, capable of delivering flexible and rapid solutions as problems arose. Centralisation and automation of trial processes with web based trial management systems were crucial. Information videos and group discussions facilitated recruitment and helped to maintain quality.

Strengths and weaknesses

In FLEXISIG, the sigmoidoscopy screening trial, the involved general practices checked the local authority lists and excluded 2% of 375 744 men and women as they were deemed ineligible owing to colorectal cancer, terminal disease, and so on.[8] In the UKCTOCS, the 3266 general practices did no cleaning of lists, as we were concerned that this would affect timely mailing of invitations, which was crucial to maintain target recruitment. Thirty two of 1.2 million women invited complained about being contacted. This included four women with ovarian cancer out of an estimated 587 who would have been invited on the basis of national incidences.[9] Given the small numbers excluded in FLEXISIG, the low rate of complaints in the UKCTOCS, and the substantial effort required, "cleaning up" of local authority lists does not seem to be necessary.

We could not account for undelivered invitations in our trial, as the initial invitations were posted in envelopes with no return address. In the recent FLEXISIG trial, 4% of invites were undelivered,[10] and this is useful in calculating accurate acceptance rates. The overall acceptance rate in the UKCTOCS was similar to that seen in the colorectal cancer screening trial at Dundee and the WISDOM trial involving oestrogen use after the menopause.[11 12] These rates were substantially higher than the 4.3% acceptance rate reported after a mass mailing of more than 3.4 million in the US systolic hypertension (SHEP) trial.[13] However, they were much lower than the 55% achieved in the UK FLEXISIG trial. One reason is the way randomisation was approached. In the UKCTOCS, we asked women to help to test a new screening programme for ovarian cancer. We told them that we would need to involve 200 000 women, half of whom would be screened and half would have the usual medical care. In FLEXISIG, participants were asked, "If you were invited to have the bowel cancer screening test, would you take up the offer?" The acceptance rate was based on those who answered "yes."[8] This emphasises the importance of wording of the invitations, an aspect sometimes overlooked in clinical trials.

In the UKCTOCS, 202 638 women were randomised in 4.3 years, which translates to a randomisation rate of 47 125 per year. The only trial to report higher rates was the FLEXISIG trial, with a randomisation rate of 68 173 per year.[10] This was achieved by adopting a two stage recruitment design; 170 483 people who agreed to attend bowel cancer screening if invited were randomised, and only 40 674 allocated

Strategies adopted to facilitate recruitment

Set-up

- Custom built trial management system to centralise and automate all trial processes such as invitation, logging of replies, scheduling of appointments, confirmation of eligibility, randomisation, and printing of letters
- Web browser and high security encryption enabling staff at all locations to enter and view data and reschedule appointments
- Specialised software commissioned from the NHS to flag women on primary care trusts' registers and allow electronic transfer of their personal and general practice details to the coordinating centre in lots of 5000 to 10 000 every quarter
- Trial website (www.ukctocs.org) for use by lay people and health professionals

Invitation

- Support of relevant patients' groups/charities
- National and local media coverage
- Mass mailing of invitations from coordinating centre
- Personal invitation with tear-off reply slip and prepaid return envelopes
- Accompanying brochure describing the goals and requirements of the trial in simple and concise terms
- Regular monitoring of acceptance rates to establish frequency and volume of mailing needed

Recruitment

- Ensuring adequate numbers of women have accepted the invitation before starting recruitment at a trial centre
- Increasing appointments in individual recruitment clinics to accommodate low attendance
- Additional large ad hoc clinics staffed by both local and coordinating centre teams
- Information video at recruitment appointment
- Interactive group discussions

Management by coordinating centre

- Senior investigator involved in day to day running of trial
- Proactive management team
- Named coordinating centre member interacting with each trial centre on a daily basis
- Prompt identification and resolution of logistical problems (staffing, delivery of consumables, information technology networking problems, postal strikes)
- Fortnightly monitoring of targets

WHAT IS ALREADY KNOWN ON THIS TOPIC

In the UK, less than one third of multicentre randomised controlled trials achieve their original recruitment target

Limited attention is paid to the management and conduct of these trials

WHAT THIS STUDY ADDS

Successful recruitment to trials needs constant monitoring by a committed proactive management team, able to rapidly deliver individual solutions as problems arise

Centralisation and automation of trial processes by use of interactive web based trial management systems are crucial in large multicentre randomised controlled trials

Recruitment can be facilitated by using information videos and group discussions

to the screen arm were recruited.[14] This design, however, does mean that only limited data and no biological samples are available in the control group. A novel feature of the UKCTOCS was the use of an information video and group discussion during recruitment. The video ensured that all participants received high quality standardised information which was sustainable. Research nurses recruiting an average of 100 women a week find it difficult to deliver the same high quality message repeatedly.

Implications of the study

Senior investigators need to set aside time to support day to day trial management. Centralisation and automation of trial processes by use of web based trial management systems with high security encryption are essential. Information videos and group discussions allow sustained delivery of high quality standardised information during prolonged recruitment. The over-riding approach needs to be one that incorporates proactive management, flexibility, and individualised solutions. Close cooperation and regular communication between the coordinating and trial centre teams are key to success.

We are particularly grateful to the women throughout the UK who are participating in the trial and to the entire medical, nursing, and administrative staff who work on the UKCTOCS.

Contributors: See bmj.com.

Funding: The trial is core funded by the Medical Research Council, Cancer Research UK, and the Department of Health with additional support from the Eve Appeal, Special Trustees of Bart's and the London, and Special Trustees of UCLH. This work was done at UCLH/UCL within the "women's health theme" of the NIHR UCLH/UCL Comprehensive Biomedical Research Centre supported by the Department of Health. SS has received research support from NCI (grant numbers CA086381 and CA083639). The researchers are independent from the funders.

Competing interests: IJ has consultancy arrangements with Vermillion and Becton Dickinson, both of which have an interest in tumour markers and ovarian cancer. They have provided consulting fees, funds for research, and staff but not directly related to this study. SS has received research support from Fujirebio Diagnostics but not in relation to this trial.

Ethical approval: The study was approved by the UK North West Multicentre Research Ethics Committees (North West MREC 00/8/34) with site specific approval from the local regional ethics committees and the Caldicott guardians (data controllers) of the primary care trusts.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 McDonald AM, Knight RC, Campbell MK, Entwistle VA, Grant AM, Cook JA, et al. What influences recruitment to randomised controlled trials? A review of trials funded by two UK funding agencies. *Trials* 2006;7:9.
2 Jacobs I, Stabile I, Bridges J, Kemsley P, Reynolds C, Grudzinskas J, et al. Multimodal approach to screening for ovarian cancer. *Lancet* 1988;i:268-71.
3 Jacobs I, Davies AP, Bridges J, Stabile I, Fay T, Lower A, et al. Prevalence screening for ovarian cancer in postmenopausal women by CA 125 measurement and ultrasonography. *BMJ* 1993;306:1030-4.
4 Jacobs IJ, Skates S, Davies AP, Woolas RP, Jeyerajah A, Weidemann P, et al. Risk of diagnosis of ovarian cancer after raised serum CA 125 concentration: a prospective cohort study. *BMJ* 1996;313:1355-8.
5 Jacobs IJ, Skates SJ, MacDonald N, Menon U, Rosenthal AN, Davies AP, et al. Screening for ovarian cancer: a pilot randomised controlled trial. *Lancet* 1999;353:1207-10.
6 Menon U, Skates SJ, Lewis S, Rosenthal AN, Rufford B, Sibley K, et al. Prospective study using the risk of ovarian cancer algorithm to screen for ovarian cancer. *J Clin Oncol* 2005;23:7919-26.
7 Menon U, Burnell M, Sharma A, Gentry-Maharaj A, Fraser L, Ryan A, et al. Decline in use of hormone therapy among postmenopausal women in the United Kingdom. *Menopause* 2007;14:462-7.
8 Atkin WS, Hart A, Edwards R, McIntyre P, Aubrey R, Wardle J, et al. Uptake, yield of neoplasia, and adverse effects of flexible sigmoidoscopy screening. *Gut* 1998;42:560-5.
9 Office for National Statistics. Cancer statistics registrations: registrations of cancer diagnosed in 2004, England. *Series MB1 35* 2004;24.
10 UK Flexible Sigmoidoscopy Screening Trial Investigators. Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial. *Lancet* 2002;359:1291-300.
11 Gray M, Pennington CR. Screening sigmoidoscopy: a randomised trial of invitation style. *Health Bull (Edinb)* 2000;58:137-40.
12 Paine BJ, Stocks NP, Maclennan AH. Seminars may increase recruitment to randomised controlled trials: lessons learned from WISDOM. *Trials* 2008;9:5.
13 Cosgrove N, Borhani NO, Bailey G, Borhani P, Levin J, Hoffmeier M, et al. Mass mailing and staff experience in a total recruitment program for a clinical trial: the SHEP experience. *Control Clin Trials* 1999;20:133-48.
14 Atkin WS, Edwards R, Wardle J, Northover JM, Sutton S, Hart AR, et al. Design of a multicentre randomised trial to evaluate flexible sigmoidoscopy in colorectal cancer screening. *J Med Screen* 2001;8:137-44.

Accepted: 5 September 2008