

## Randomised comparison of three methods of administering a screening questionnaire to elderly people: findings from the MRC trial of the assessment and management of older people in the community Topic: 145;9;10;35

Liam Smeeth, Astrid E Fletcher, Susan Stirling, Maria Nunes, Elizabeth Breeze, Edmond Ng, Christopher J Bulpitt, Dee Jones

### Abstract

**Objective** To compare three different methods of administering a brief screening questionnaire to elderly people: post, interview by lay interviewer, and interview by nurse.

**Design** Randomised comparison of methods within a cluster randomised trial.

**Setting** 106 general practices in the United Kingdom.

**Participants** 32 990 people aged 75 years or over registered with participating practices.

**Main outcome measures** Response rates, proportion of missing values, prevalence of self reported morbidity, and sensitivity and specificity of self reported measures by method of administration of questionnaire for four domains.

**Results** The response rate was higher for the postal questionnaire than for the two interview methods combined (83.5% *v* 74.9%; difference 8.5%, 95% confidence interval 4.4% to 12.7%,  $P < 0.001$ ). The proportion of missing or invalid responses was low overall (mean 2.1%) but was greater for the postal method than for the interview methods combined (4.1% *v* 0.9%; difference 3.2%, 2.7% to 3.6%,  $P < 0.001$ ). With a few exceptions, levels of self reported morbidity were lower in the interview groups, particularly for interviews by nurses. The sensitivity of the self reported measures was lower in the nurse interview group for three out of four domains, but 95% confidence intervals for the estimates overlapped. Specificity of the self reported measures varied little by method of administration.

**Conclusions** Postal questionnaires were associated with higher response rates but also higher proportions of missing values than were interview methods. Lower estimates of self reported morbidity were obtained with the nurse interview method and to a lesser extent with the lay interview method than with postal questionnaires.

### Introduction

Multidimensional assessment was originally developed in response to studies finding high levels of undetected

medical or social problems among elderly people,<sup>1 2</sup> showing the need for a systematic approach to detection of problems. The effectiveness of multi-dimensional assessment for elderly people has been assessed in several randomised trials, with mixed results, and considerable uncertainty exists about the optimal method of administering an initial screening questionnaire.<sup>3 4</sup> Self administered postal questionnaires have several advantages over face to face interviews.<sup>5</sup> They are cheaper, large numbers can be completed more quickly, interviewers do not have to be recruited and trained, and high rates of response have been obtained among elderly people.<sup>6-8</sup>

Since 1990, primary care teams in England and Wales have been required to offer an annual screening assessment to all patients aged 75 and over (the "over 75 check").<sup>9</sup> Although the contract for general practitioners in England and Wales recommends that people aged 75 and over are invited to "participate in a consultation," a postal approach has been advocated for initial screening.<sup>10-13</sup> The national service framework for older people includes a recommendation that elderly people should receive some form of single assessment "which is matched to their individual circumstances."<sup>14</sup> The consultation papers for the single assessment process state that either postal or face to face methods may be appropriate.<sup>15</sup>

We present a randomised comparison of three different methods of administering a screening questionnaire: post, interview by lay interviewer, and interview by nurse. All three methods have previously been used in randomised controlled trials of multidimensional screening assessments of elderly people, but none of these trials compared different methods.<sup>16-18</sup> We compared rates of response, proportions of missing data, estimates of prevalence, and the validity of the results obtained. The data presented are from the baseline assessments of the Medical Research Council trial of the assessment and management of older people in the community.

Centre for Ageing and Public Health, London School of Hygiene and Tropical Medicine, London WC1E 7HT

Liam Smeeth  
*clinical lecturer in epidemiology*

Astrid E Fletcher  
*professor of epidemiology and ageing*

Susan Stirling  
*research fellow*

Elizabeth Breeze  
*lecturer*

Edmond Ng  
*research fellow*

Section of Care of the Elderly, Imperial College, Faculty of Medicine, Hammersmith Campus, London W12 0NN

Maria Nunes  
*research officer*

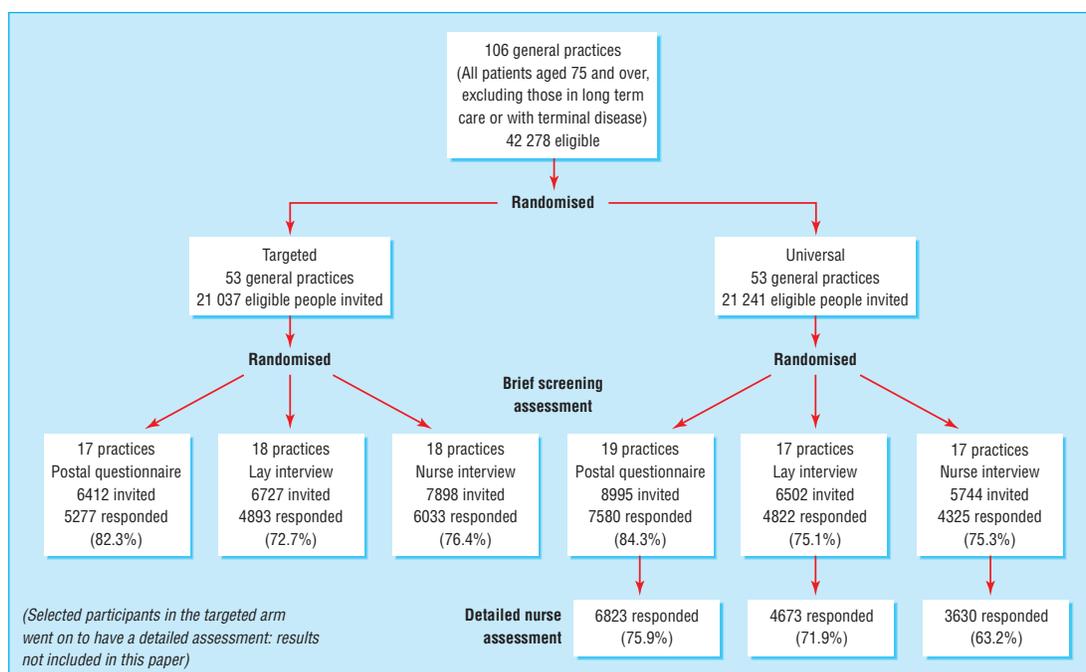
Christopher J Bulpitt  
*professor of geriatric medicine*

Research Team for Care of Elderly People, University Department of Geriatric Medicine, University of Wales College of Medicine, Cardiff CF64 2XX

Dee Jones  
*principal research fellow*

Correspondence to: L. Smeeth  
liam.smeeth@lshtm.ac.uk

BMJ 2001;323:1-7



Numbers randomised, interventions offered, and response rates

## Methods

### The MRC trial

The MRC trial of the assessment and management of older people in the community is a community based randomised controlled trial comparing different approaches to multidimensional screening for people aged 75 years and over. The trial was designed to recruit 108 practices with an average of 500 eligible patients per practice, stratified to provide a representative sample in terms of mortality (standardised mortality ratio) and deprivation (Jarman score) of general practices in the United Kingdom. All practices were part of the MRC general practice research framework. In each practice, all patients aged 75 years or over registered with the practice were included in the study, unless they were resident in a long stay hospital or nursing home or were terminally ill.

We randomised practices to two groups: targeted screening and universal screening. We invited all participants to have a brief screening assessment. All participants in the universal screening arm were then invited to have a more detailed assessment. In the targeted screening arm, only people found to have a pre-specified level of problems during the brief assessment were invited to have a detailed assessment. The figure shows the study design, number of eligible participants, and rates of response.

A statistician drew up a computer generated randomisation list, stratified by tertiles of Jarman score and standardised mortality ratio; practices were randomly allocated centrally at the London School of Hygiene and Tropical Medicine as they were recruited to the trial. Because of the nature of the intervention, participants and researchers could not be blinded to the group assignment. One hundred and six practices participated in the study. The original target was 108 practices, with 18 in each of the six different combinations of universal or targeted and post, lay interview, or

nurse interview (see fig). The sample sizes were based on the main trial outcomes (mortality and admissions to hospital or institution).

One practice in the universal-post group split up after randomisation, and the participating general practitioner consequently had a smaller number of patients. For this reason we recruited a further practice to that group, giving a total of 19 practices. Three practices withdrew at a late stage (after randomisation but before data collection) and were not replaced: hence three of the six groups include only 17 practices.

Data collection took place between 1995 and 1999. Before starting the assessments, the nurses and lay interviewers attended a training session. A few practices joined the study late, in which case the training took place at the practice. The nurses involved in the study were mostly practice nurses involved in practice based research; some of the nurses devoted all their time to research.

We obtained ethical approval for all aspects of the study from the relevant ethics committees.

### Assessment of participants

#### Brief assessment

We invited all participants in both arms of the study to have a brief screening assessment consisting of a range of health related questions covering the areas specified in the 1990 contract. We randomised practices to one of three methods of administering the brief assessment: postal questionnaire, interview by lay interviewer, or interview by nurse. We used 26 screening questions plus questions about smoking, alcohol intake, and physical activity included largely for purposes of epidemiological research. We present results for these three items here because they are specifically recommended by the national service framework for older people for inclusion in a screening assessment.<sup>14</sup> Most of the questions had a graded

**Table 1** Rates of response by method of administration. Values are percentages (numbers) unless stated otherwise

Age (years) and sex	Method of administration			
	Post	Lay interviewer	Nurse interviewer	All methods
Men:				
75-79	87.9 (2587)	79.2 (1911)	79.0 (1976)	82.4 (6474/7860)
80-84	83.7 (1399)	76.0 (1121)	76.6 (1204)	78.9 (3724/4719)
85-89	82.9 (651)	75.4 (491)	79.9 (543)	79.6 (1685/2116)
≥90	73.0 (154)	67.2 (125)	74.7 (133)	71.7 (412/575)
All	85.4 (4791/5610)	77.2 (3648/4727)	78.2 (3856/4933)	80.5 (12 295/15 270)
Women:				
75-79	85.1 (3598)	74.3 (2661)	76.6 (2845)	79.0 (9104/11 523)
80-84	82.1 (2564)	71.8 (1984)	73.3 (2014)	76.0 (6562/8634)
85-89	80.5 (1359)	70.2 (1044)	72.7 (1142)	74.7 (3545/4747)
≥90	72.0 (545)	66.2 (438)	74.0 (501)	70.8 (1484/2096)
All	82.3 (8066/9796)	72.1 (6127/8497)	74.7 (6502/8707)	76.7 (20 695/27 000)
All:				
% (95% CI)	83.5 (80.9 to 85.7)	73.9 (69.5 to 77.8)	75.9 (70.4 to 80.7)	78.0 (75.6 to 80.6)
No	12 857/15 407	9775/13 229	10 358/13 642	32 990/42 278*

\*Overall total differs slightly from totals for men plus women because data on sex were missing for 8 participants.

response. For example, possible answers to questions about hearing and vision were “no difficulty,” “a little difficulty,” and “a lot of difficulty.”

#### Detailed assessment

We then invited all participants in the universal screening arm to have a more comprehensive detailed assessment by the trained nurse. In the targeted screening arm, only participants with a predetermined number and type of problems at the brief assessment went on to have the detailed assessment. We do not present the results of the detailed assessments in the targeted arm of the trial here because the participants were not a representative group of people. For four of the domains included in the brief assessment, more accurate and thorough assessments were undertaken in the detailed assessment.

**Hearing**—We gave participants the whispered voice test, in which specified numbers and letters are spoken in a whisper at full expiration by a tester standing 15 cm behind the patient. The test was performed with participants wearing hearing aids if they used them, thus testing participants’ everyday hearing. The whispered voice test has been found to have a sensitivity of between 80% and 100% and a specificity of between 80% and 89% when compared with measurement of hearing loss in the range 30-40 decibels by pure tone audiometry.<sup>19-23</sup>

**Vision**—We measured participants’ distance visual acuity at 3 m with a Glasgow acuity chart.<sup>24</sup> For this study, we compared the screening question about vision with a binocular visual acuity cut-off point equivalent to a Snellen acuity of less than 6/18. Interventions such as cataract extraction would usually be offered when a visual acuity of less than 6/18 is found,<sup>25</sup> and this level of visual impairment is below the level required to be legally permitted to drive in the United Kingdom.<sup>26</sup>

**Depression**—Participants completed the 15 item version of the geriatric depression scale.<sup>27-28</sup> All questions are answered yes or no. Using a cut-off score of ≥6 to indicate depression, the 15 item version of the geriatric depression scale has been found to have a sensitivity in the range 78-85% and a specificity in the range 74-82% when compared with the Diagnostic and Statistical Manual of Mental Disorders<sup>29</sup> or Geriatric Mental

State<sup>30</sup> criteria.<sup>31-33</sup> We used a cut-off score of ≥6 in this study.

**Cognition**—Participants completed the mini-mental state examination.<sup>34</sup> This examination is a widely used test of cognitive function and has been shown to be both valid and reliable.<sup>34-36</sup> It has two sections: a verbal section with a maximum score of 21 and a performance section (involving, for example, copying a drawing) with a maximum score of 9. For physical or educational reasons, not all people are able to complete the performance section. The nurse administering the questionnaire decided whether participants were able to complete the performance section. Cut-off points of less than 17 for the whole test or less than 12 for the verbal section were used in the main trial to indicate the need for referral and were used to indicate likely cognitive impairment in this study.

#### Analyses

All data were recorded on specially designed forms and scanned electronically. We analysed the data with Stata 6 software. Because we used stratified cluster sampling by general practice, additional variance could have arisen because observations on individuals within clusters may be correlated (that is, individuals within the same practice may be more similar than individuals in different practices). All analyses took account of the cluster design in the estimation of standard errors.<sup>37-39</sup> We analysed all participants with available data in the groups to which they were initially randomised.

Because of the large number of possible comparisons, we performed hypothesis testing only for comparisons of particular interest. For example, in looking at the sensitivity and specificity of self reported measures by method of questionnaire administration there are 48 possible significance tests. We present confidence intervals to aid interpretation.

## Results

#### Response rates

From 42 278 eligible participants, we obtained data from the brief screening assessment for 32 990 people, an overall response rate of 78.0%. Men were more likely to respond than women (80.5% *v* 76.7%,  $P < 0.001$ ), and this sex difference persisted after adjustment for age: the adjusted odds ratio for

**Table 2** Proportion of missing or invalid responses and prevalence of self reported problems by method of administration of questionnaire. Values are percentages\*

Screening domain	Missing values				Reported prevalence of problem†			
	Post (n=12 857)	Lay interviewer (n=9775)	Nurse interviewer (n=10 358)	All (n=32 990)	Post (n=12 857)	Lay interviewer (n=9775)	Nurse interviewer (n=10 358)	All (n=32 990)
No personal help	3.4	0.8	1.4	2.0	7.3	4.8	3.2	5.3
Rarely see people	3.1	0.4	0.5	1.5	6.5	4.1	2.7	4.6
More than 7 drugs‡	6.4	0.4	1.9	3.2	5.1	4.8	7.1	5.6
Feel sad or depressed	2.3	0.9	1.5	1.7	8.5	9.6	8.8	8.9
A lot of difficulty seeing	2.0	0.7	0.9	1.3	11.0	10.5	7.8	9.8
A lot of difficulty hearing	2.1	0.2	0.5	1.0	8.5	9.4	5.3	7.8
Problems with memory	2.0	0.3	0.7	1.1	10.4	9.4	7.9	9.3
Difficulty with activity and no help available:								
Cutting toe nails	3.8	0.4	1.1	2.0	7.4	2.9	2.6	4.5
Dressing self	2.7	0.2	0.8	1.4	0.5	0.1	<0.1	0.3
Cooking hot meal	4.6	1.2	2.7	3.0	1.9	0.2	<0.1	0.8
Housework	4.0	1.0	2.3	2.6	3.1	0.6	0.2	1.4
Stairs	7.9	2.6	3.4	4.9	2.9	1.6	0.7	1.8
Washing	2.7	0.3	0.9	1.5	2.0	1.5	1.3	1.6
Walking 50 yards	4.5	1.6	2.5	3.0	3.5	2.2	1.1	2.3
Shopping	4.2	1.2	2.2	2.7	3.4	0.5	0.2	1.5
Keeping home warm	2.8	0.5	0.6	1.4	1.5	1.4	0.5	1.1
Making ends meet	3.8	0.6	1.4	2.1	0.8	1.0	0.3	0.7
Managing finances	3.6	1.4	2.0	2.4	3.6	1.8	0.7	2.1
Recent history of:								
Vomiting blood	6.7	0.3	0.7	2.9	0.4	0.2	0.1	0.3
Coughing up blood	7.6	0.4	0.7	3.3	0.6	0.7	0.4	0.5
Shortness of breath at rest	8.1	0.5	0.9	3.6	10.7	8.6	3.9	7.8
Shortness of breath when talking	10.2	0.5	1.1	4.5	13.2	9.0	5.0	9.3
Swelling of legs	8.2	0.6	0.9	3.7	13.6	7.0	3.6	8.3
Unexpected weight loss	7.3	1.0	1.6	3.6	7.9	5.4	3.2	5.6
Wetting self‡ (>1 per week)	3.5	0.4	0.7	1.7	15.5	12.4	10.8	13.1
Soiling self‡ (>1 per week)	2.7	0.5	1.1	1.5	3.3	1.5	1.3	2.1
Falls‡ (>4 in past 6 months)	2.0	0.4	0.3	1.1	1.9	1.9	1.1	1.6
Additional questions:								
Not at all physically active	1.5	0.2	0.3	0.7	6.0	4.9	4.6	5.2
Currently smoke	1.4	0.3	0.7	0.8	9.3	9.0	9.4	9.2
Currently drink alcohol	2.8	0.32	1.12	1.5	62.3	71.2	72.7	68.3

\*Numbers are available from the author.

†Omitting missing values.

‡Included a supplementary question to quantify the problem.

response comparing men with women was 1.22 (95% confidence interval 1.16 to 1.29,  $P < 0.001$ ). Data on sex were missing for eight eligible participants and could not be clarified as they had moved away. Responders were slightly younger than non-responders (median 80.3 years *v* 81.0 years,  $P < 0.001$ ). Table 1 shows rates of response by method of administration of the brief screening assessment broken down by age and sex.

The response rates were 83.5% for the postal method, 73.9% for the lay interview method, and 75.9% for the nurse interview method. The mean response rate for the two interview methods was 74.9%, a difference from the postal method of 8.5% (4.4% to 12.7%,  $P < 0.001$ ). The response rate for the nurse interview method was 2% higher (-4.4% to 8.5%,  $P = 0.53$ ) than that for the lay interview method.

The response rate fell with increasing age in both men and women. In addition, the higher response rate seen for the postal approach was not apparent among the oldest age groups. This may have been because of the higher levels of home visiting undertaken by the nurse or lay interviewer in the oldest age groups. For both nurse interview and lay interview methods, around 30% of assessments were undertaken in people's homes (31.2% and 30.0%,  $P = 0.78$ ). The pro-

portion of assessments undertaken in people's own homes rose from around 16% in the 75-80 year age group to around 75% in the 90 years and over age group, with no significant differences between the two interview methods.

#### Proportion of missing or invalid responses and prevalence of problems

Table 2 shows the proportion of missing or invalid responses and the prevalence of problems reported. The proportions of missing or invalid responses were higher in the postal questionnaire group than in the interview groups for all questions. The mean proportion of missing responses was 2.1% overall: 4.1% for postal questionnaire, 0.6% for lay interview, and 1.1% for nurse interview. The mean proportion of missing responses was 3.2% higher (2.7% to 3.6%,  $P < 0.001$ ) for the postal approach than for the two interview methods combined. The difference between the two interview methods was small: 0.5% higher (0.12 to 0.93,  $P = 0.017$ ) for the nurse interview group.

Four of the questions (as indicated in table 2) were in two parts: the first part determined whether a domain applied, and this was followed by a branching question to quantify the problem. For example, the first

**Table 3** Sensitivity and specificity of self reported measures in four domains by method of administration (see text for criteria used). Values are percentages (95% CI)

Method of administration	Vision		Hearing		Depression		Cognition	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Post	40.2 (34.8 to 46.0)	94.2 (92.8 to 95.3)	23.2 (18.8 to 28.4)	95.2 (93.0 to 96.7)	40.3 (33.9 to 47.0)	93.2 (91.5 to 94.6)	51.0 (38.8 to 63.1)	91.0 (89.8 to 91.9)
Lay interviewer	38.7 (31.3 to 46.7)	93.4 (91.4 to 95.0)	30.2 (25.3 to 35.6)	95.0 (93.0 to 96.4)	47.9 (43.5 to 52.4)	91.5 (89.1 to 93.3)	42.7 (31.8 to 54.3)	92.4 (90.8 to 93.8)
Nurse interviewer	26.2 (19.9 to 33.6)	95.6 (93.7 to 97.0)	20.8 (15.0 to 26.6)	97.4 (96.2 to 98.3)	40.5 (34.2 to 47.0)	93.2 (91.4 to 94.6)	34.8 (28.9 to 41.2)	93.0 (91.2 to 94.5)
All methods	36.4 (32.7 to 40.3)	94.3 (93.4 to 95.1)	25.0 (22.0 to 28.2)	95.7 (94.6 to 96.6)	42.8 (39.3 to 46.4)	92.6 (91.5 to 93.6)	43.8 (37.4 to 50.4)	92.0 (91.1 to 92.6)

question about incontinence asked if there was ever a problem and was followed by a question to determine frequency. For these domains, a missing response refers to one or both parts of the question being missing. The patterns of missing responses for these four domains were similar to the patterns for the other domains.

In the postal questionnaire group, 21% of responders stated that someone had helped them to fill in the questionnaire. The mean proportion of missing responses was slightly lower for people who had received help filling in the questionnaire (3.8% *v* 4.4%,  $P = 0.03$ ).

For 22 out of the 26 screening domains the prevalence of self reported problems was higher for the postal group than for the interview groups (the exceptions were medications, hearing, depression, and financial difficulty). Some of the observed differences were small, but larger differences were seen for many domains. To a much lesser degree, a similar pattern was seen in the lay interview group and the nurse interview group, with slightly lower levels of problems reported in the nurse interview group across all but two domains. For the three additional questions that were not originally intended for screening purposes (physical activity, smoking, and alcohol intake) no clear differences were seen in the answers obtained by the three methods.

#### Detailed assessment

All eligible people in the universal screening arm of the trial were invited to have a more detailed assessment by a nurse. Of 21 241 people, 15 126 (71.2%) responded; 64% of detailed assessments took place in general practice surgeries and 36% in people's homes. The median time between the brief assessment and the detailed assessment was 13 days (interquartile range 7 to 21 days).

#### Sensitivity and specificity of self reported measures

We compared four of the domains included in the brief assessment with more accurate and objective assessments undertaken in the detailed assessment. Table 3 shows the sensitivity and specificity for the self reported measures. The sensitivity varied by method of administration, with a lower sensitivity for the nurse interview method for three of the four domains assessed (depression was the exception). However, the 95% confidence intervals for the three different methods overlapped within each domain. The sensitivity of the questions was generally low (all less than 51%). The specificity of the self reported measures was high for all domains, with little variation by method of questionnaire administration.

## Discussion

The main strengths of this study are its randomised design and the fact that it is the largest study of its kind yet undertaken in this age group. Our results are directly relevant to the "single assessment" described in the national service framework for older people.<sup>15</sup> The response rate was substantially higher for the postal questionnaire than for either interview method. Previous studies have found higher response rates,<sup>8</sup> no difference,<sup>40-41</sup> or lower response rates for postal questionnaires.<sup>42-44</sup> It is likely that people's ability or willingness to travel to the general practice affected the response rates in the interview groups. It is also possible that people prefer filling in a questionnaire.

Men were more likely to respond than women for all methods. As far as we are aware this is a novel finding. However, with the exception of one study that found no sex difference in response rates,<sup>45</sup> previous studies of screening older people have not analysed sex differences of responders and non-responders.

The proportion of missing or invalid responses in this study was low overall, but the proportion of missing responses was significantly greater for the postal method, in line with previous work.<sup>41-43-46</sup> Two previous studies have found that for intimate or sensitive issues the proportions of missing data were higher for the interview method.<sup>8-40</sup> We did not see this phenomenon in this study.

Although the reliability (that is, repeatability) of postal questionnaires and interview administered questionnaires has been compared and generally been found to be high,<sup>6-8</sup> the validity of the information obtained by the two approaches is less well established. In a study of outcomes after surgery, participants had a tendency to give more positive or optimistic answers in a self completed questionnaire than in an interview.<sup>8</sup> In studies of alcohol dependence, participants were less likely to report excessive drinking and less likely to report adverse symptoms in self administered questionnaires than in interviews.<sup>47-48</sup> However, in another study people were less likely to give a socially undesirable response to an interviewer than in a postal questionnaire.<sup>49</sup> In a household health survey in which individual responses were compared with health records, the accuracy of responses was higher for postal questionnaires.<sup>40</sup>

The levels of self reported morbidity in our study were substantially lower in the two interview groups than in the postal questionnaire group. In addition, slightly lower levels of morbidity were reported to nurse interviewers than to lay interviewers. These findings are difficult to explain with any certainty. The

response rate in the postal group was higher. The higher levels of morbidity observed in the postal group could have occurred if people with higher levels of morbidity were more likely to respond to a postal questionnaire than to an interview assessment. The higher proportions of missing values for the postal approach could also have inflated estimates of morbidity if people who did not have a problem were more likely to miss out a question because it did not seem relevant to them. Another possibility is that people were more likely to report potentially embarrassing problems, such as incontinence, in a postal questionnaire than in a face to face interview. However, the patterns of both missing data and reported prevalence were the same for potentially sensitive or embarrassing questions as for questions unlikely to cause embarrassment.

In the four domains for which it was possible to assess the diagnostic accuracy of the self reported measures used in the brief assessment (visual acuity, whispered voice test, geriatric depression scale, and mini-mental state examination), specificity was high and similar for the three methods across all four domains. The sensitivity was somewhat lower for the nurse interviewer group for three of the domains. The explanation for this finding is not clear. The relatively low sensitivities of simple questions (by any of the three methods of administration) for the detection of poor vision, hearing impairment, depression, and cognitive problems suggest that caution is needed in the use of simple questions for case finding.

In everyday practice the rates of response, proportions of missing data, and validity of the answers are likely to depend on the particular questions being asked and the skills of the interviewer as well as on the method of administration. The nurse interviewers used in this study were usually the practice nurses, and the lay interviewers were often clerical staff in the practice. Both groups received identical training in administering the brief questionnaire. They were fairly typical practice staff, not highly trained researchers, and this represents a feasible use of practice staff. The use of a postal questionnaire is supported by the higher response rate and because it is likely to be much cheaper than using an interviewer. The only clear disadvantage of the postal technique in this study was a higher proportion of missing or invalid responses, but even this higher level was only around 4%.

Whether the differences observed for the three methods affect health outcomes (mortality, hospital admission rates, and quality of life) and cost effectiveness will be answered by the ongoing randomised trial.

We thank the nurses, general practitioners, other staff, and patients in the participating practices; everyone at the MRC general practice research framework coordinating centre, particularly Jeannett Martin and Nicky Fasey; clerical staff at London School of Hygiene and Tropical Medicine (Janbibi Mazar and Rakhi Kabawala) and Hammersmith Hospital (Ruth Peters) for all their work on the study; Amina Latif (research officer) and Elaine Springer (clerical officer) of the University of Wales College of Medicine for administrative assistance; Judith Nickson (University of Cambridge) and Jennifer Evans and Richard Wormald (Moorfields Eye Hospital) for advice and help with training the research nurses; Alistair Tulloch (University of Oxford) for advice; and the trial steering committee—J Grimley Evans (chair), A Haines (previous chair), K Luker, C Brayne, M Vickers, M Drummond, S Lonsdale, and L Davies.

## What is already known on this topic

The optimum method of administering a brief multidimensional screening assessment to elderly people is not known

## What this study adds

Postal questionnaires produce a higher response rate than interviews by nurses or lay interviewers but also higher proportions of missing data

Interview by nurses and to a lesser degree by lay interviewers is associated with lower levels of self reported morbidity than are postal questionnaires

Contributors: LS had the idea for the study, analysed the data, and wrote the paper. AEF is the principal investigator and CJB and DJ are co-investigators of the MRC trial of the assessment and management of older people in the community and designed and implemented the trial. SS devised and carried out the randomisation procedure, monitored data collection, and took part in training the nurses and lay interviewers. MN and EB were involved in administering the study and editing data. EN took part in data management and cleaning. All authors commented on drafts of the paper. LS and AEF are the guarantors.

Funding: The MRC trial of the assessment and management of older people in the community was funded by the UK Medical Research Council, the Department of Health, and the Scottish Office. LS is funded by a research fellowship from London NHS Executive.

Competing interests: None declared.

- 1 Anderson WF, Cowan NR. A consultative health centre for older people: the Rutherglen experiment. *Lancet* 1955;ii:239-40.
- 2 Williamson J, Stokoe IH, Gray S, Fisher M, Smith A, McGhee A, et al. Old people at home: their unreported needs. *Lancet* 1964;i:1117-20.
- 3 Fletcher A. Multidimensional assessment of elderly people in the community. *Br Med Bull* 1998;54:945-60.
- 4 Van Haastregt JC, Diederiks JP, van Rossum E, de Witte LP, Crebolder HF. Effects of preventive home visits to elderly people living in the community: systematic review. *BMJ* 2000;320:754-8.
- 5 Bowling A. Data collection methods in quantitative research: questionnaires, interviews and their response rates. In: *Research methods in health*. Buckingham: Open University Press, 1997:228-49.
- 6 McDonnell H, Long AE, Harrison BJ, Oldman C. A study of persons aged 65 and over in the Leeds metropolitan district. *J Epidemiol Community Health* 1979;33:203-9.
- 7 Wallace P, Haines A. Use of a questionnaire in general practice to increase the recognition of patients with excessive alcohol consumption. *BMJ* 1985;290:1949-53.
- 8 Doll H, McPherson K, Davies J, Flood A, Smith J, Williams G, et al. Reliability of questionnaire responses as compared with interview in the elderly: views of the outcome of transurethral resection of the prostate. *Soc Sci Med* 1991;33:1303-8.
- 9 Department of Health and the Welsh Office. *General practice in the National Health Service. A new contract*. London: Department of Health, 1989.
- 10 Barber JH, Wallis JB, McKeating E. A postal screening questionnaire in preventive geriatric care. *J R Coll Gen Pract* 1980;30:49-51.
- 11 Wilcock GK. Use of a self-administered postal questionnaire when screening for health problems in the elderly. *Gerontology* 1979;25:345-9.
- 12 Victor CR. Some methodological aspects of using postal questionnaires with the elderly. *Arch Gerontol Geriatr* 1988;7:163-72.
- 13 Bowns I, Challis D, Tong MS. Case finding in elderly people: validation of a postal questionnaire. *Br J Gen Pract* 1991;41:100-4.
- 14 Department of Health. *National service framework for older people*. London: Stationery Office, 2001.
- 15 Department of Health. The single assessment process: consultation papers and process. [www.doh.gov.uk/scg/sap](http://www.doh.gov.uk/scg/sap) (accessed 16 Aug 2001).
- 16 Carpenter GI, Demopoulos GR. Screening the elderly in the community: controlled trial of dependency surveillance using a questionnaire administered by volunteers. *BMJ* 1990;300:1253-6.
- 17 Pathy MS, Bayer A, Harding K, Dibble A. Randomised trial of case finding and surveillance of elderly people at home. *Lancet* 1992;340:890-3.
- 18 Wagner EH, LaCroix AZ, Grothaus L, Leveille SG, Hecht JA, Artz K, et al. Preventing disability and falls in older adults: a randomised controlled trial. *Am J Public Health* 1994;84:1800-6.
- 19 Swan IR, Browning GG. The whispered voice as a screening test for hearing impairment. *J R Coll Gen Pract* 1985;35:197.
- 20 Macphree GJ, Crowther JA, McAlpine CH. A simple screening test for hearing impairment in elderly patients. *Age Ageing* 1988;17:347-51.
- 21 Uhlmann RF, Rees TS, Psaty BM, Duckert LG. Validity and reliability of auditory screening tests in demented and non-demented older adults. *J Gen Intern Med* 1989;4:90-6.

- 22 Browning GG, Swan IR, Chew KK. Clinical role of informal tests of hearing. *J Laryngol Otol* 1989;103:7-11.
- 23 Eekhof JA, de Bock GH, de Laat JA, Dap R, Schaapveld K, Springer MP. The whispered voice: the best test for screening for hearing impairment in general practice? *Br J Gen Pract* 1996;46:473-4.
- 24 McGraw P, Winn B. Glasgow acuity cards: a new test for the measurement of letter acuity in children. *Ophthalmic Physiol Opt* 1993;13:400-4.
- 25 Frost NA, Sparrow JM. Use of vision tests in clinical decision making about cataract surgery: results of a national survey. *Br J Ophthalmol* 2000;84:432-4.
- 26 Drivers Medical Group. *At a glance guide to current medical standards of fitness to drive*. Swansea: DVLA, 2000.
- 27 Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1982;17:37-49.
- 28 Sheik JL, Yesavage JA. Geriatric depression scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol* 1986;37:819-20.
- 29 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders: DSM-IV*. Washington, DC: APA, 1994.
- 30 Copeland JR, Dewey ME, Griffiths-Jones HM. A computerized psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGECAT. *Psychol Med* 1986;16:89-99.
- 31 D'Ath P, Katona P, Mullan E, Evans S, Katona C. Screening, detection and management of depression in elderly primary care attenders. I: The acceptability and performance of the 15 item geriatric depression scale (GDS15) and the development of short versions. *Fam Pract* 1994;11:260-6.
- 32 Herrmann N, Mitmann N, Silver IL, Shulman KI, Busto UA, Shear NH. A validation of the geriatric depression scale short form. *Int J Geriatr Psychiatry* 1996;11:457-60.
- 33 Almeida OP, Almeida SA. Short versions of the geriatric depression scale: a study of their validity for the diagnosis of a major depressive episode according to ICD-10 and DSM-IV. *Int J Geriatr Psychiatry* 1999;14:858-65.
- 34 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.
- 35 Nelson A, Fogel BS, Faust D. Bedside cognitive screening instruments. A critical assessment. *J Nerv Ment Dis* 1986;174:73-83.
- 36 Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc* 1992;40:922-35.
- 37 Cochran W. *Sampling techniques*. New York: John Wiley, 1977.
- 38 Rao JN, Scott AJ. The analysis of categorical data from complex sample surveys: chi-squared tests for goodness of fit and independence in two-way tables. *J Am Statistical Assoc* 1981;76:221-30.
- 39 Rao JN, Scott AJ. On chi-squared tests for multiway contingency tables with cell proportions estimated from survey data. *Annals Stats* 1984;12:46-60.
- 40 Siemiatycki J. A comparison of mail, telephone, and home interview strategies for household health surveys. *Am J Public Health* 1979;69:238-45.
- 41 Addington-Hall J, Walker L, Jones C, Karlens S, McCarthy M. A randomised controlled trial of postal versus interviewer administration of a questionnaire measuring satisfaction with, and use of, services received in the year before death. *J Epidemiol Community Health* 1998;52:802-7.
- 42 Cartwright A. Interviews or postal questionnaires? Comparisons of data about women's experiences with maternity services. *Milbank Q* 1988;66:172-89.
- 43 Mallinson S. The short-form 36 and older people: some problems encountered when using postal administration. *J Epidemiol Community Health* 1998;52:324-8.
- 44 Picavet HSJ. National health surveys by mail or home interview: effects on response. *J Epidemiol Community Health* 2001;55:408-13.
- 45 Hanger HC, Sainsbury R. Screening the elderly: a Christchurch study. *NZ Med J* 1990;103:473-5.
- 46 Hayes V, Morris J, Wolfe C, Morgan M. The SF-36 health survey questionnaire: is it suitable for use with older adults? *Age Ageing* 1995;24:120-5.
- 47 Cutler SF, Wallace PG, Haines AP. Assessing alcohol consumption in general practice patients—a comparison between questionnaire and interview (findings of the Medical Research Council's general practice research framework study on lifestyle and health). *Alcohol Alcohol* 1988;23:441-50.
- 48 Wiseman EJ, Heithoff KA. Comparison of DSM-III-R symptoms for alcohol dependence between patient self-report and clinician interview or the structured clinical interview for DSM-III-R. *J Addict Dis* 1996;15:43-54.
- 49 Pless IB, Miller JR. Apparent validity of alternative survey methods. *J Community Health* 1979;5:22-7.

(Accepted 27 September 2001)