

Development of a modified Winchester disability scale—the elderly at risk rating scale

Ian P Donald

Abstract

Objective—To show that the elderly at risk rating scale (EARRS) satisfies the requirements of an assessment tool for routine health checks in people over 75 and would also be suitable as a method of collecting epidemiological data on the needs of the elderly in a locality.

Design—Development and validation of a questionnaire based on a modification of the Winchester rating scale, by a series of prospective, comparative studies before the use of the instrument in a community survey.

Setting—Elderly care day hospital and the community.

Subjects—Elderly patients referred to an elderly care day hospital; population survey of subjects over 75 living at home.

Main outcome measures—Reliability of responses using the kappa statistic; comparison of the scale with the Barthel index of daily living.

Results—EARRS has satisfactory validity and reliability when repeated by the same observer or a different observer, with a mean weighted kappa score above 0.80 in both instances. As a measure of disability in the community, it is better than the Barthel score in that it avoids the ceiling effect. The score is correlated with age, social situation, and receipt of support services, and individual questions scale appropriately to adverse outcomes.

Conclusion—The EARRS satisfies the requirements of an assessment tool for health checks in the elderly. It is suitable for both population surveys and routine practice in primary care, has proved popular with practice nurses, and is easy to complete.

(*J Epidemiol Community Health* 1997;51:558-563)

The introduction of a compulsory annual health check for people over the age of 75 in the new contract for general practitioners of 1990¹ should lead to appropriate documentation that can both assist the interviewer in systematically covering the areas mentioned in the contract and provide a useful record for subsequent reference or analysis. No questionnaires or standardised forms were recommended by the NHS Executive when the check up was introduced, however, and the development of an instrument that could be used locally and could save GPs from having to devise their own seemed opportune. It was also tempting to explore

whether such a questionnaire could also be used to measure disability, since functional independence is the focus of most of the areas that have to be covered in the health check. Furthermore, if such an instrument were adopted universally within an area it would provide an easy way of collecting information about the elderly and their needs in that locality. I describe the development of such an instrument in this article.

The starting point for the instrument was the Winchester disability rating scale.² This scale was designed for the purpose of surveillance of the elderly in the community. For each area of assessment there is a choice of five descriptive categories of competence. This range was attractive because it was felt that a "yes/no" scale would be too crude to detect the early stages of failing independence at home. The detection of small yet important changes on a five point scale might offer the potential for intervention to prevent further deterioration and dependency. The Winchester scale includes assessment in 19 areas, with 15 of these ranked on a score of 1-5. It has been subjected to limited validation.³ The use of medicines was not covered by the scale, however, and mobility did not encompass outdoor activity so modification seemed to be required.

Method

A small working party was formed in 1989, comprising a consultant physician (IPD), a general practitioner (JRJ), a consultant in public health (PS), and a community nurse (MR), with advice from a consultant psychiatrist (JL). This group discussed the wording of questions before and after a trial of the questionnaire in hospital and in the community. The new instrument was required to assess the areas specified in the 1990 contract for general practitioners (section 13D, subparagraph 5) for elderly health checks.¹ These areas were as follows:

- Sensory function,
- Mobility,
- Mental condition,
- Physical condition,
- Continence,
- Social environment, and
- Use of medicines.

Our working party felt that the assessment of mobility needed to encompass the full range, from driving or using public transport to confinement to bed. To do this adequately, the mobility question needed to be expanded to cover mobility indoors and outdoors. The assessment of mental health was expanded so

Elderly Care Unit,
Gloucestershire Royal
Hospital, Great
Western Road,
Gloucester GL1 3NN
I P Donald

Correspondence to:
Dr I P Donald.

Accepted for publication
January 1997

that there were separate questions on depression, confusion, and bereavement. The Winchester question on "health" was rephrased "energy", as this was considered to be a better subjective way of assessing the sense of wellbeing. New questions on diet, joint pains, use of (and comprehension of) medicines, and the need for chiropody were included. The question on recent falls was also rephrased so that it could be included in the overall scoring system. This resulted in a new scale comprising 20 questions rating each area in five categories and an additional two questions without scoring on the type of residence and domestic social support. The instrument, called the elderly at risk rating scale (EARRS) is presented in table 1. It has four sections. The first is demographic and includes name, address, date of birth, categories of residence, and six support services, to be highlighted when applicable. The remaining sections cover physical risk, mental risk, and social risk. The first has 12 questions and the other two sections have four questions each. These 20 questions are all ordinal data, with a range of five possible answers to each question. These answers are always arranged hierarchically, and scored 1 to 5, where 1 describes lowest risk or optimal independence and 5 describes greatest risk or highest dependency.

A prototype version of the questionnaire was piloted by three doctors and three nurses in 40 patients all over the age of 70, living at home, and including a wide range of age and independence. Each patient was interviewed on two occasions in their own home, by a doctor and a nurse. Verbal consent for all interviews was obtained by the district nurse or GP over the telephone, at which time a convenient appointment was arranged. Inter-

KEY POINTS

- The elderly at risk rating scale (EARRS) is a brief yet comprehensive assessment of the elderly at home and satisfies the Department of Health's requirements for health checks in people over 75 years.
- The scale is easily completed by a nurse in 10 minutes.
- EARRS has excellent reproducibility enabling comparisons between current and previous assessments.
- It is suitable for collecting epidemiological data on functional disability.

views were within a week of each other. The working party met together again after this pilot exercise and discovered that there had been differing interpretations of particular questions. Some questions were rephrased, one was changed completely, and instructions, to be printed on the reverse were rewritten.

The modified version of the questionnaire (table 1) was examined in 20 subjects who were attending an elderly care day hospital on their first occasion for assessment. They were rated one week apart, but on this occasion by the same nurse interviewer. The nurses were not connected with the group who had designed the questionnaire, nor had they participated in the earlier pilot study. They were instructed to perform the second test without having the first form available, and were informed that the purpose was to see if the patient's responses changed from week to week.

Comparisons between the EARRS, Barthel,⁴ and the Nottingham extended activities of daily living instrument (ADL)⁵ were made with 52

Table 1 The elderly at risk rating scale

Surname:	Forename:	Action required:			
Address:	BP:	Weight:	Urine:		
Date of birth:	Flat	Bungalow	House	Warden-aided	Institution
Type of residence ?					
Support service ?	Private help	Meals on wheels	District nurse	Day centre	Day hospital
Home help	1	2	3	4	5
Physical risk	Drives/public transport	Walks daily	Walks occasionally	Walks to gate	Housebound
1 Outdoor mobility	Independent	Uses stick around house	Uses frame around house	Room-fast	Bed/chairfast
2 Indoor mobility	0	1	2	3	More
3 Falls in past 3 months	Bath/shower independent	Bath shower with difficulty	Can strip wash	Bath/shower with help	Needs washing
4 Washing	Independent	Some difficulty	Much difficulty	Manages with help	Needs dressing
5 Dressing	Independent	Commode/bottle at night	Commode/bottle day&night	Occasional accidents	Frequent accidents
6 Toilet	Independent, good diet	Independent, limited diet	Independent, poor diet	Eats with difficulty	Needs feeding
7 Eating / diet	Satisfactory	Slight impairment	Satisfactory with aid	Hard of hearing, no aid	Deaf +/- aid
8 Hearing	Satisfactory	Some difficulty	Cannot read	Cannot watch TV	Blind
9 Sight	None	Occasional	Some of the time	A lot of the time	All of the time
10 Joint pains	No problem	Mild difficulty	Receives foot care	Receives chiropody	Needs but gets no chiropody
11 Feet/chiropody	None prescribed	1-3 drugs, understands	>3 drugs, understands	Little vague	Erratic / no idea
12 Drugs	1	2	3	4	5
Mental risk	Boundless	Average	Limited	Very limited	Almost none
13 Energy	None	—	Close relative last year	—	Close relative last 3 months
14 Bereavement	Not at all	Occasionally	Some of the time	A lot of the time	All of the time
15 Sad/weepy	Normal	Occasionally	Forgetful	Muddled	Very confused
16 Confusion	1	2	3	4	5
Social risk	Relatives/other	—	Spouse only	—	Alone
17 Lives with?	None required	Less than daily	Daily	Twice daily	Continuous
18 All help received	None required	Carers have no difficulty	Carers have some difficulty	Carers under stress	Carers cannot continue
19 Carers					Very poor
20 Housing suitability	Ideal	Adequate	Inadequate/problems	Untidy/hazardous	

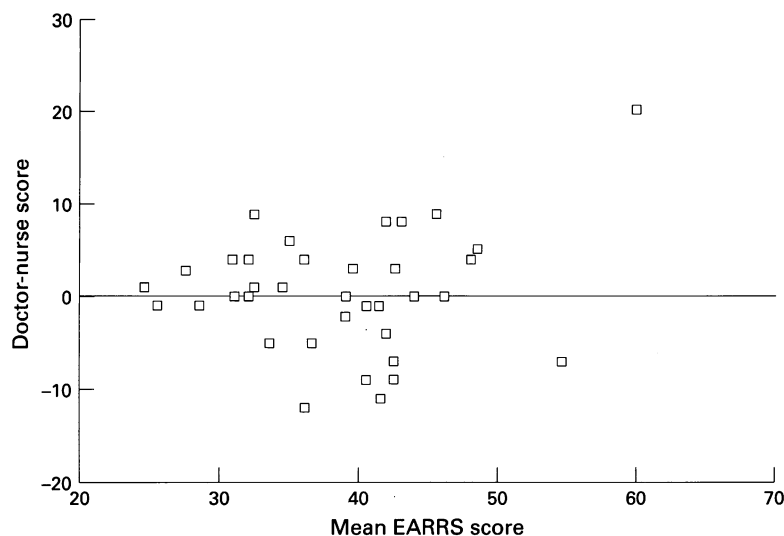


Figure 1 Comparison of doctors' and nurses' scores using the elderly at risk rating scale (EARRS).

Table 2 Results of elderly at risk rating scale administered by a doctor and nurse on separate occasions to 40 elderly subjects at home

Observer	Mean	Median	Range (SD)
Nurse	38.2	39	24-58 (7.8)
Doctor	38.7	38	25-70 (8.8)

elderly people attending two elderly care day hospitals and living at home. These subjects also had the EARRS and Barthel scales repeated at the end of their treatment within the day hospital. All assessments were carried out by nursing staff.

The EARRS was proposed by the local health authority in 1990 as a model for conducting the health checks in the over 75s. Twenty five out of 89 practices made some use of the scale. Seven of these practices were asked to participate in a longitudinal study of ageing and disability, and EARRS scores from the first year of this study are also presented. Analyses have shown that the survey respondents, who comprised 61% of the over 75 population, were similar to the non-responders in terms of age, sex, use of health service resources, and subsequent mortality.

Limits of agreement for the mean versus difference Bland-Altman plot have been defined,⁶ and were used here. Statistical analysis of the reliability of responses in the questionnaire was performed using the kappa statistic with quadratic weighting. Quadratic weighting enables the kappa coefficient to be applied to several categories, with the coefficient reflecting the level of disagreement between the two interviewers—eg how frequently the difference was only one category. Interpretation of the kappa result is taken from Altman.⁷ Paired *t* tests with two tailed significance were used to compare two assessments of the same patients; unpaired *t* tests were used in the analysis of data from the longitudinal study of ageing. Comparisons between EARRS and other instruments used Spearman's rho rank order correlation coefficient, because the Barthel data were considerably skewed. Tests were run on SPSS for Windows Release 6.1.2.

Table 3 Interobserver reliability of elderly at risk rating scale

No	Weighted kappa	No	Weighted kappa
1	0.66	11	0.84
2	0.76	12	0.71
3	0.96	13	0.90
4	0.71	14	0.80
5	0.75	15	0.85
6	0.90	16	0.84
7	0.78	17	0.98
8	0.88	18	0.74
9	0.74	19	0.84
10	0.84	20	0.87

Results

FUNCTIONALITY

The layout of the scale on one side of a folded A4 card was acceptable to nurses and doctors, while a paper version may be easier to file in old-style primary health care records if this is required. It was easy to complete the form with a highlighter pen and this method was adopted by all users; no free text entries were required. The arrangement of all questions in a similar manner made scoring very simple, and also enabled a rapid assessment of the patient's profile at a glance, identifying the main problem area. For example, highlighting towards the right hand side of the form in the social risk section or the physical risk section indicated principal problems in these areas.

The time taken to complete the scale was determined by the style of interview and the extent to which the interviewer explored problems raised by the instrument. Typically, the instrument is completed in 10-15 minutes. The detail of the instrument struck a reasonable balance between speed and a worthwhile and comprehensive assessment.

INTEROBSERVER RELIABILITY (NURSE-DOCTOR COMPARISON)

The nurse scores were similar to the doctor scores, with an identical overall mean (table 2). However, this covered some large individual differences in scores, which ranged from 0-20. The median difference was 4 points. The Bland-Altman plot is shown in figure 1 and emphasises the occurrences of discrepant scoring, with the "limits of agreement",⁶ incorporating 2SDs of the differences; these were -11.6 to +13.6 points. In 23 subjects, the doctor conducted the first interview, in 11 the nurse conducted the first interview, and in 6 the two interviews were carried out on the same day. The mean score for the first interview did not differ from the mean score of the second interview (38.8 *v* 38.6, *p*>0.1, paired *t* test).

Agreement on individual questions was considered more important than the overall score. This was examined using the kappa coefficient with quadratic weightings; the results are shown in table 3.

TEST/RETEST RELIABILITY (SAME NURSE OBSERVER)

The EARRS scores from these repeated interviews by the same observer one week apart showed a mean difference of 1 point, giving 95% confidence intervals (CI) of agreement on the Bland Altman plot of -2.5, +3.4 points.

Table 4 Test-retest reliability of elderly at risk rating scale

No	Weighted kappa	No	Weighted kappa
1	0.94	11	0.99
2	0.89	12	0.95
3	0.92	13	0.88
4	0.94	14	1.0
5	0.94	15	0.79
6	0.94	16	0.75
7	0.92	17	0.75
8	0.87	18	0.80
9	0.90	19	0.89
10	0.80	20	0.93

This suggests that a clinically important change in the EARRS can only be considered when there has been a change greater than 5.9 points⁶: a change of 6 points may be taken as indicating a significant change. This marked improvement on the interobserver results reflects not only the expected greater reliability of the score when repeated by the same observer, but also the improvement with the modified scale and the benefit of some experience with the use of the instrument.

The kappa tests for individual questions were very satisfactory, showing good or very good agreement throughout. The mean weighted kappa was 0.81, with all results >0.75. The kappa figures are shown in table 4. For the 16 questions unchanged from the interobserver study above, the kappa values were significantly improved using the same observer (paired *t* test, $p < 0.001$).

COMPARISON OF THE EARRS WITH OTHER INSTRUMENTS

The EARRS has been compared with both the Barthel index and the Nottingham extended ADL. It is apparent on looking at these scales that they are very different in the areas covered. The Barthel index reflects the first 10 questions of the EARRS, but does not include assessment of drugs, foot care, mental health, and social support. The Nottingham extended ADL includes questions on activity in the kitchen and other domestic tasks, as well as leisure activities; none of these areas are covered by the EARRS except for one question on preparation of meals.

Box plots for the Barthel index (transformed to scale from 20 to 100) and EARRS are shown in figure 2, which illustrates the greater ceiling effect of the Barthel score in elderly subjects at home rather than in hospital. A scatter plot is given in figure 3; and the correlation coefficient was -0.65 ($p < 0.0001$), as the EARRS scales in the opposite direction to the Barthel.

The EARRS has also been compared with the extended Barthel index in 20 subjects attending an elderly care day hospital, but the correlation was poor; $r = -0.5$ ($p = 0.02$).

SENSITIVITY TO CHANGE

This has been observed in the context of an elderly care day hospital treatment programme. While these treatment programmes focus on rehabilitation and reducing disability, it has been difficult to demonstrate improvement on global scales, because these improvements are often marginal and limited to resolving specific problems.⁸ In a group of 52 patients

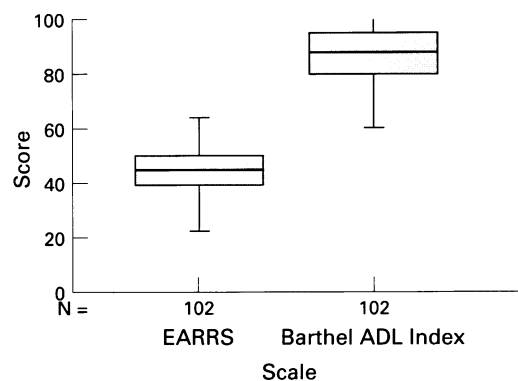


Figure 2 Elderly at risk rating scale (EARRS) compared with the Barthel index scores in 102 elderly people assessed at home.

completing treatment, the EARRS score changed by a median of -1 and a mean of -2.0 , with a range of -14 to $+9$. By comparison, in the same patients, the Barthel score changed by a median of 1 , a mean of 1.0 and a range of -4 to $+6$. There was no change in Barthel score for 19 patients and in the EARRS for 15 patients. If a change in the Barthel score of more than 2 points, and a change in the EARRS greater than 5 points (as described above) is considered significant, each instrument described change in 14 patients, with agreement in 12. A scatter plot correlating the two changes in score is shown in figure 4; the correlation is $r = 0.27$, $p = 0.047$.

EARRS SCORES IN FIRST YEAR OF AN OVER 75 LONGITUDINAL STUDY

In the longitudinal study, mean scores show a progressive rise with age, and are higher in females than males (fig 5). The type of residence was not significantly associated with the score, except for institutions, where the mean score was 45.8 for residents compared with 30 for non-residents. Those living with their spouse had the lowest mean score of 28.3; those living with their family had the highest mean score of 34.9; those living alone had intermediate scores, with a mean of 31.4 (unpaired *t* tests, all differences significant, $p > 0.0001$). The scores for those in receipt of support services are shown in table 5. All questions were found to scale correctly, with an increased score being associated with a greater risk of adverse outcomes (details will be reported in future publication). The exception was question 17, "Lives with?", where living alone was not associated with subsequent morbidity or mortality.

Discussion

This new scale has good face validity and is easy for an unaccustomed user to understand; indeed the original Winchester scale was designed for use by volunteers.^{2,9} It covers a large number of the areas traditionally covered when any member of the health care team is conducting a home based interview with an older patient. It is quick to complete, facilitated by the use of a highlighter pen, and the scoring is straightforward and does not require any calculator. The scale has been designed to include

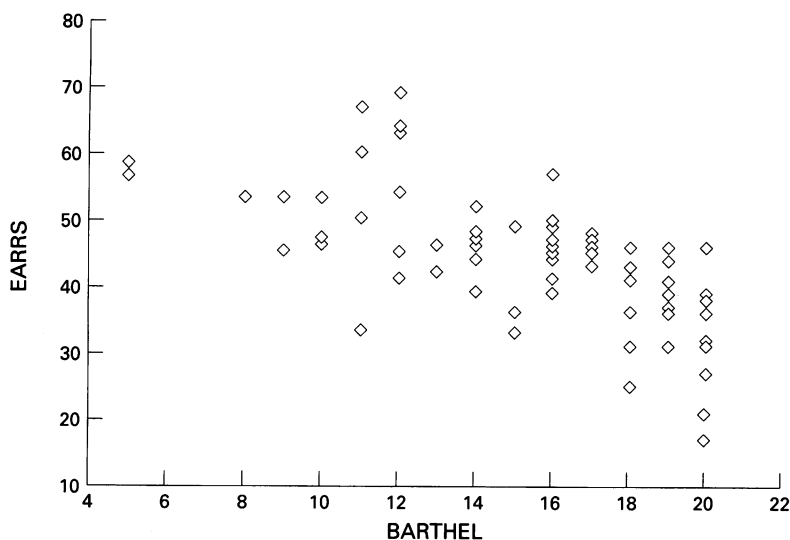


Figure 3 Elderly at risk rating scale (EARRS) compared with the Barthel index scores in elderly people assessed in an elderly care day hospital.

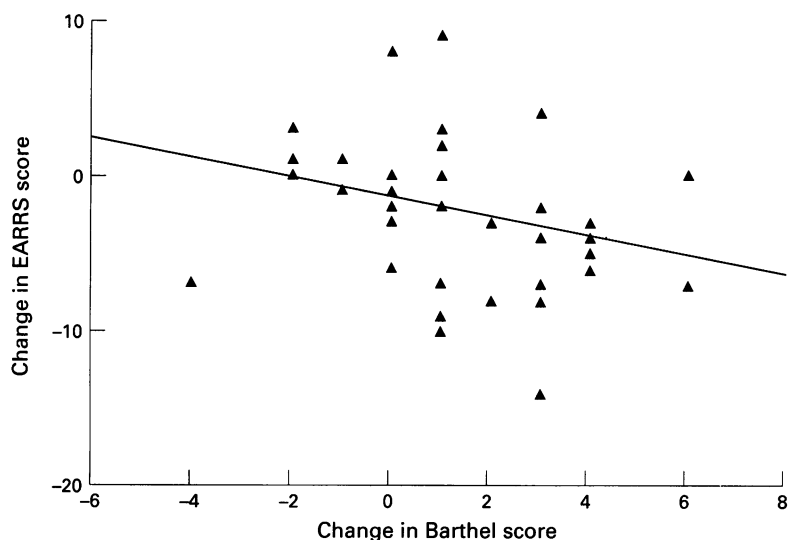


Figure 4 Changes after treatment programmes in subjects assessed in elderly care day hospitals - elderly at risk rating scale (EARRS) versus the Barthel index.

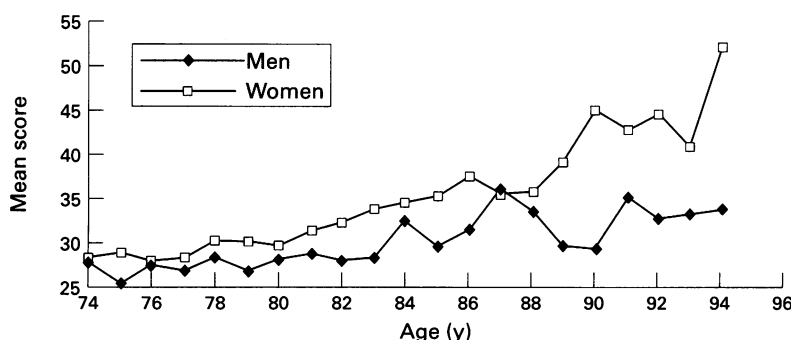


Figure 5 Mean elderly at risk rating scale scores in relation to age and sex in a community survey of 1815 subjects over 75 years of age.

all the areas specified within the contractual requirements for GPs when conducting the annual health check on people over 75.

The EARRS was the most popular scale adopted locally by GPs for their routine over 75 health checks. It was acceptable to practice nurses, and at the same time enabled opportunistic collection of longitudinal data on ageing in the community. It was seen to be a helpful tool in

Table 5 Mean elderly at risk rating scale scores for those with and without support services

Service	Those in receipt of support service		Not receiving support
	% of sample	Mean score	Mean score
Home help	13	37.6	30.3
Private help	9	34.4	30.9
Meals on wheels	4	40.1	30.9
District nurse	8	42.3	30.3
Day centre	3	40.3	30.9
Day hospital	1	44.5	31.1

undertaking the health check, rather than a burden to complete.

The reliability of the instrument is also very important. The doctor-nurse interobserver study showed some potential difficulties, while, as would be expected, the scores were more reliable when repeated by the same observer. The mean weighted kappa score between different observers was still very acceptable at 0.82, and we have found that greater familiarity with the wording of the scales has reduced discrepant scoring between different observers. The addition of simple instructions on the reverse side of the sheet on which the scale is printed scale, focused on clarifying the areas of uncertainty, has also reduced any ambiguity. The reliability has been measured only in selected samples of elderly persons, which limits the generalisability of the findings. However the test-retest study simulated the most common application for the EARRS, namely where the same nurse is repeating the assessment at a later date.

An expected degree of correlation has been shown with the widely accepted Barthel scale; a closer level of agreement would have been surprising because the conventional Barthel scale does not assess ability within the home. The EARRS has been shown to be better than the Barthel instrument because it does not have a ceiling effect when used with elderly subjects living at home.

The ability of the EARRS score to change during the course of a rehabilitation programme has been shown, although the changes in most individuals were modest. This probably reflects the true situation that indeed little change has taken place in some patients. Change was recorded more frequently by the EARRS than by the Barthel instrument, but the occurrence of significant change was similar.

The validity of EARRS is further indicated by the increase in the score with age, and the relationship between response and adverse outcome for 19 out of 20 questions. This important area will be the subject of future reports. However, these correlations support the use of EARRS in the prospective collection of epidemiological data which would be useful in planning future service needs. This is an important additional benefit from routine health checks, which has been recommended by health care planners,¹⁰ but rarely adopted by general practitioners.

A comparable tool, the functional autonomy measurement system (SMAF), has been developed in Canada.¹¹ This comprises 29 items

apart from details of residence, companion, and support and is therefore more thorough than the EARRS. The EARRS has comparable reliability to SMAF, and its advantage in Britain must be that it has been developed and tested within the UK, has been edited to a reasonable minimum to satisfy the Department of Health guidelines, and is sufficiently easy to use that it can be administered by volunteers. Comparisons of the SMAF with the EARRS or the Winchester scale have not been carried out. The more widely used Barthel and other ADL scales are less suited to a global assessment of an older person living at home, because they have been designed to measure the progress of a stroke patient in hospital⁴ or at home.⁵

The modest changes made to the Winchester scale coupled with the further validation of this version of the instrument makes it ideal in fulfilling the purpose of a short yet comprehensive assessment tool for health checks in the elderly.

I am grateful to Dr John Roccyn-Jones, Dr Janet Lugg, Dr Peter Strangeways, and Sister Maria Rainer for their assistance in the development of this scale, and to Gloucestershire FHSA for their enthusiasm in supporting the use of the instrument within the county. I am also grateful to Professor C Bulpitt, Royal Postgraduate Medical School, Hammersmith Hospital, London, for considerable help and advice in the writing of this manuscript.

Funding: Gloucestershire Health Authority Research and Development.

Conflicts of interest: none.

- 1 Department of Health. *Terms of service for doctors in general practice*. London: Department of Health, November 1989.
- 2 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.
- 3 Oliveri S, Carpenter GI, Demopoulos GR. Validity and reliability of the Winchester disability rating scale(2). *Gerontology* 1994;40:319-24.
- 4 Collin C, Wade DT, Davies S, Horne V. The Barthel ADL index: a reliability study. *International Disability Studies* 1988;10:61-3.
- 5 Nouri F, Lincoln NB. An extended activities of daily living scale for stroke patients. *Clinical Rehabilitation* 1987;1:301-5.
- 6 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;i:307-10.
- 7 Altman DG. *Practical statistics for medical research*. London: Chapman and Hall, 1991.
- 8 Donald IP, Baldwin RN, Bannerjee M. Gloucester hospital-at-home: a randomized controlled trial. *Age Ageing* 1995;24:434-9.
- 9 Beales DL. *The use of trained volunteers in a screening programme: an evaluative study*. Occasional Paper 35. London: Royal College of General Practitioners, March 1987.
- 10 Bone MR, Bebbington AC, Jagger C, Morgan K, Nicolaas G. *Health expectancy and its uses*. London: HMSO, 1995;73-7.
- 11 Hébert R, Carrier R, Bilodeau A. The functional autonomy measurement system (SMAF): Description and validation of an instrument for the measurement of handicaps. *Age Ageing* 1988;17:293-302.