

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

PENNY D'ATH,* PHILIPPA KATONA,** ELEANOR MULLAN,*† SANDRA EVANS‡ AND CORNELIUS KATONA*†

D'Ath P, Katona P, Mullan E, Evans S and 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. *Family Practice* 1994; 11: 260-266.

One-hundred and ninety-eight elderly subjects attending their general practitioners (GPs) were asked to complete the 15 item Geriatric Depression Scale (GDS15). Analysable results were obtained from 194 (98%). Of these, 67 (34%) scored above the GDS15 cut-off (4/5) for significant depressive symptomatology. 87.6% found the questionnaire to be acceptable and only 3.6% found it very difficult or very stressful. The GDS15 had a high level of internal consistency (Cronbach's alpha = 0.80). All the individual items of the GDS15 associated significantly ($P < 0.01$) with total score and 'caseness'. A single question "do you feel that your life is empty?" identified 84% of 'cases'. In an attempt to devise short scales to screen elderly primary care patients for depression, the data were subjected to logistic regression analysis. Ten (GDS10), four (GDS4) and one (GDS1) item versions were generated. Agreement between these short scales and the GDS15 in the original sample was 95, 91 and 79% respectively. Cronbach's alpha was 0.72 for the GDS10 and 0.55 for the GDS4. The short scales were then validated in an independent sample of 120 patients in whom both GDS data and the results of a detailed psychiatric interview (the Geriatric Mental Status Schedule, GMS) were available. The sensitivity and specificity of the GDS10 against GMS caseness were 87 and 77% (cut-off 3/4); those of the GDS4 were 89 and 65% (cut-off 0/1) and 61 and 81% (cut-off 1/2). Sensitivity and specificity for the GDS1 were 59 and 75%. It is concluded that these short scales may be useful in helping GPs and practice staff to identify elderly patients with significant depressive symptoms.

INTRODUCTION

Depression in old age carries a poor prognosis¹ with increased use of health and social service facilities² and

excess mortality³ as well as frequent chronicity of the depressive symptoms.⁴ However, while depressive symptoms are common in elderly patients attending their general practitioners (GPs),^{5,6} depression is seldom recorded in the GP case notes even where a clear diagnosis can be made.⁷ It is also clear that depression in old age, as identified through community surveys, is seldom treated.^{4,8} These findings suggest that depression in old age may be both underdetected and undertreated in primary care.

* Psychiatric Wing, Princess Alexandra Hospital, Hamstel Road, Harlow, Essex, ** Lower Clapton Health Centre, Lower Clapton Road, London E5, † Department of Psychiatry, University College London Medical School, Wolfson Building, Middlesex Hospital, London W1N 8AA and ‡ St Bartholomew's Hospital Medical School, UK. Correspondence to C.K. at UCLMS.

In a younger population, recognition of depression by the GP improves outcome even when the patient does not comply with treatment.⁹ Validating a patient's feelings by making a formal diagnosis may itself have therapeutic potential. The possible contribution of a short questionnaire, acceptable to both GPs and their patients, that would improve detection rates and facilitate a higher rate of effective intervention, is thus considerable. Screening for such depressive symptoms would be a legitimate part of the mandatory screening of patients aged 75 and over, which now forms part of the GP contract.¹⁰

The Geriatric Depression Scale (GDS)¹¹ is a 30 item yes/no questionnaire devised specifically to detect depression in elderly subjects. It has been extensively validated in hospital samples.¹² Evans and Katona⁶ have shown it to have acceptable sensitivity (85%) and specificity (68%) against diagnosis based on structured psychiatric interview (the Geriatric Mental Status Schedule (GMS))¹³ in a British primary care sample aged 65 and over. It also agreed with GMS diagnosis significantly more often than GPs' own diagnoses (76% versus 65%). A shorter, 15 item version of the Geriatric Depression Scale (GDS15) has been devised by Sheikh and Yesavage.¹⁴ The GDS15 has been recommended for use within the over-75 health check by the Royal College of General Practitioners.¹⁵

The first aim of the present study was to assess the feasibility and acceptability of using the GDS15 to detect depression in elderly GP surgery attenders. Since the GDS15 takes up to 10 minutes to complete, it remains relatively cumbersome to be an ideal part of the routine screening of elderly patients (within a brief consultation with the GP, or in the context of a brief domiciliary health screen administered by non-medical practice staff). A second aim of the study was to devise shorter and more acceptable versions of the GDS.

METHODS

The study was performed at the Lower Clapton Health Centre, a seven partner general practice that serves a population of approximately 10 500 patients. The practice has a computerized appointment system linked to its age-sex register, enabling easy identification of all elderly surgery attenders. All patients aged 65 or over attending their GPs during weekday morning, afternoon or evening surgeries over an 8 week period between January and March 1993 were approached in the waiting room by a researcher (PD) and asked if they minded participating in a study examining whether a short questionnaire was useful in helping doctors to know how their patients are feeling in terms of mood.

If there were no objections to participating, PD then asked the 15 questions of the GDS15. For 10 of the questions the answer 'yes' gives a positive score (indicating depression); in the remaining five the answer 'no' scores positively. The scores are then summed to

give a total of 0-15, with a score of 5 or more indicating probable depression.

In addition, as a measure of acceptability of the questionnaire, patients were asked three questions: "did you find the questionnaire difficult?", "did you find the questionnaire stressful?" and "did you find the questionnaire acceptable?". These were scored as "not at all", "quite" or "very".

Patients who required home visits or who attended weekend surgeries were excluded from data collection but weekday emergency surgery attenders were included. The weekday emergency clinic runs parallel to the other clinics and is for patients who need to see a doctor that day and are unable to make an appointment due to all clinic lists being full.

Student's *t*-test and the Wilcoxon test were used to examine the relationship between individual item scores and the GDS15 total and the internal consistency of the scale examined using Cronbach's alpha. In addition, subjects were divided into cases and non-cases on the basis of a GDS15 cut-off of 4/5. The relationship between individual item scores and caseness/non-caseness was examined using the chi-square test, as was the relationship between caseness and test acceptability. Subjects' scores on the individual GDS15 items were subjected to logistic regression analysis, with GDS15 'caseness' as the dependent variable. At each step in the analysis, the most discriminating item was eliminated and the remainder reanalysed. The procedure was continued until no further items significantly improved the goodness of fit of the logistic regression model. Ten, four and one item short versions (GDS10, GDS4 and GDS1) were derived; these are described in the Results below. The internal consistencies of the GDS10 and GDS4 were assessed using Cronbach's alpha, and correlations (Pearson's and Kendall's) were calculated between scores on the GDS15 and each of the short GDSs.

The short versions of the GDS were then tested in an independent validation sample. This consisted of the first 120 patients from the dataset collected by Evans and Katona⁶ in whom GMS diagnoses were available and individual GDS30 (and thus GDS15) items had been entered. Their demographic characteristics, derived GDS15 scores and proportion identified as being 'cases' of depression on the GMS are summarized in Table 1. Scores on the GDS10, GDS4 and GDS1 were calculated for each subject in the validation sample. Correlations, between-group comparisons and chi-square tests were calculated as for the internal validation. In addition, sensitivity (percentage GMS cases detected by each short GDS), specificity (percentage of short GDS detected cases confirmed by GMS) and negative predictive value (percentage of short GDS non-cases confirmed as non-cases by GMS) were calculated. All data analysis was carried out using SPSS/PC+, version 3.1.¹⁶

RESULTS

Demographic Characteristics of Patients Studied

As shown in Table 1, the mean age of the subjects was 73.5 years. Eighty-six subjects (43%) were aged 75 or over. The female:male ratio was approximately 2:1.

GDS Total Scores

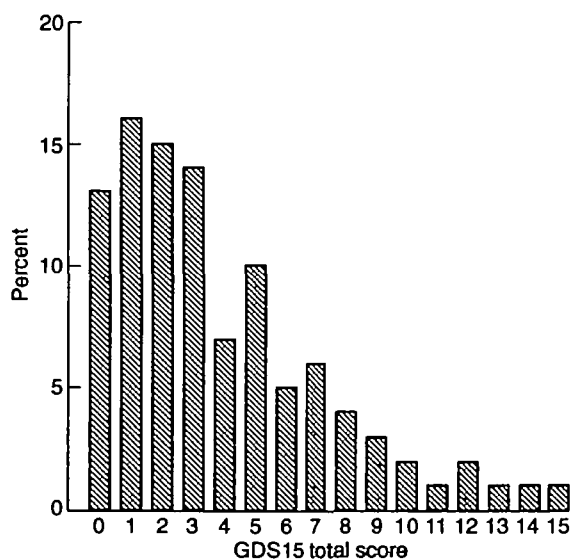
The mean total score for the 15 questions was 3.7 (median 3, range 0–15). The distribution of GDS15 total scores is shown in Figure 1. Table 1 summarizes the proportion of patients scoring above the 4/5 cut-off. There were no statistically significant relationships between GDS15 score and age (caseness rate of 33% in those aged 65–74 and 35% in those aged 75 and over). Similarly, there was no relationship between the GDS15 score and gender, with 35% of the men and 33% of the women scoring in the depressed range.

Relationship Between Individual Item Scores and GDS15

For each of the GDS15 questions, there was a statistically significant relationship ($P < 0.01$) between

TABLE 1 *Demographic and psychiatric characteristics*

Characteristic	Derivation sample	Validation sample
<i>n</i>	198	120
Sex (F/M)	126/72	82/38
Mean age (range)	73.5 (65–92)	74.1 (65–90)
Mean GDS15 score (range)	3.7 (0–15)	4.7 (0–13)
% GDS15 cases	34	48
% GMS cases	N/A	38

FIGURE 1 *The distribution of GDS15 total scores*

a positive score and the GDS15 total. Similarly significant associations were found between each item and caseness/non-caseness. The proportion of subjects giving 'depressed' responses to each GDS question (in the total group and after division into GDS15 'cases' and 'non-cases') is summarized in Table 2. Question 3 (do you feel that your life is empty?) was found to be the best predictor of caseness with an 84% level of agreement (chi-square = 76.2, $P < 0.0001$) while question 9 (do you prefer to stay at home, rather than going out and doing new things?), had the most modest statistical relationship (chi-square = 8.6, $P < 0.01$) with caseness. Cronbach's alpha for the GDS15 was 0.80.

Acceptability of the GDS15

Of 198 patients approached, only four refused to participate, leaving 194 (98%) for subsequent analysis. A total of 88% found the questionnaire to be 'very' acceptable, 89% found it to be 'not at all' stressful and 88% found it 'not at all' difficult (Table 3). There were no statistically significant relationships between gender or age and any of the parameters of acceptability. As can be seen from Table 3, there was a statistically significant relationship between GDS15 'caseness' and finding the questionnaire 'difficult' or 'stressful', with a similar trend linking 'caseness' and the experience of the GDS15 as unacceptable.

Logistic Regression Analysis

The results of the pooled within-groups correlations between individual items of the GDS15 are summarized in Table 4, showing the order of elimination and residual correlation coefficients. A 10 item version, GDS10, was derived by totalling the scores of all the items contributing significantly to the logistic regression model. Similarly, a four item version, GDS4, was derived from the totals for the four best discriminating items. The questions forming GDS10 are shown in Table 5, with the GDS4 items in bold.

Internal Validation of the Short GDSs

The GDS10 and GDS4 scores both correlated very highly with the GDS15 ($r = 0.97$ and $r = 0.89$ respectively, both $P < 0.0001$). Student's *t*-test on GDS15 scores in those scoring yes versus no on the single most discriminating question (GDS1) gave $t = 12.6$, $P < 0.0001$. Cronbach's alpha was 0.72 for the GDS10 and 0.55 for the GDS4. Using cut-off scores of 3/4 for the GDS10 and 1/2 for the GDS4 gave 93 and 91% agreement with GDS15; chi-square was 129.8 and 115.2 respectively, both $P < 0.0001$. Using lower cut-offs of 2/3 (GDS10) and 0/1 (GDS4) gave 95 and 81% agreement with GDS15. With these cut-off scores, chi-square was 144.1 and 77.5, again both $P < 0.0001$. Agreement between GDS1 'yes' and GDS15 caseness was 84%; chi-square = 76.2; $P < 0.0001$. Results for non-parametric tests (Kendall's tau and Wilcoxon's test) gave similar results, all with significance levels of $P < 0.0001$.

TABLE 2 Relationships between individual item scores and GDS15

Item	Per cent scoring + ve			Per cent agreement
	Cases (67)	Non-cases (127)	Total (194)	
Are you basically satisfied with your life?	49	5	20	79
Have you dropped many of your activities and interests?	55	19	31	72
Do you feel that your life is empty?	70	9	30	84
Do you often get bored?	78	22	41	78
Are you in good spirits most of the time?	43	8	20	75
Are you afraid that something bad is going to happen to you?	48	13	25	73
Do you feel happy most of the time?	45	3	18	79
Do you often feel helpless?	49	8	22	77
Do you prefer to stay at home, rather than going out and doing new things?	34	15	22	68
Do you feel you have more problems with memory than most?	30	7	15	71
Do you think it is wonderful to be alive now?	36	9	18	72
Do you feel pretty worthless the way you are now?	43	4	18	78
Do you feel full of energy?	73	39	51	65
Do you feel that your situation is hopeless?	42	6	18	76
Do you think that most people are better off than you are?	49	14	26	73

TABLE 3 Acceptability of the GDS15

	GDS cases (n)	GDS non-cases (n)	Total	
			n	Per cent
Difficult				
Not at all	52	118	170	87.6
Quite	9	8	17	8.8
Very	6	1	7	3.6
Stressful				
Not at all	49	123	172	88.7
Quite	12	3	15	7.7
Very	6	1	7	3.6
Acceptable				
Not at all	3	1	4	2.1
Quite	9	11	20	10.3
Very	55	115	170	87.6

Independent Sample Validation

The product-moment correlation between the GDS10 and the full GDS30 was 0.91 ($P < 0.0001$); for the 4 item GDS it was 0.76 ($P < 0.0001$). For the GDS1, comparison between GDS30 scores for subjects scoring yes and no gave $t = 7.3$, $P < 0.0001$. The corresponding non-parametric tests gave very similar results, in all cases also significant at the 0.0001 level. Results for agreement against the GMS interview for the GDS10 and GDS4 (using both possible cut-off points in each case) and for GDS1 are summarized in Table 6. The performance of the GDS15 at cut-off points of 4/5 and 5/6 (also derived from GDS30 data) in the validation sample is also shown for comparison. Since the GDS4 had higher sensitivity at the lower (0/1) and higher specificity at the higher (1/2) cut-off, we also examined the discriminating ability of the GDS10 in the 36 subjects with an intermediate score (1) on the GDS4. The GDS10 showed 74% sensitivity and 58% specificity in this subsample (chi-square = 4.4; $P < 0.05$).

TABLE 4 *Logistic regression analysis*

GDS question	Order of elimination	Correlation
1. Are you basically satisfied with your life?	3	0.29
2. Have you dropped many of your activities and interests?	8	0.25
3. Do you feel that your life is empty?	1	0.54
4. Do you often get bored?	N/A	NS
5. Are you in good spirits most of the time?	N/A	NS
6. Are you afraid that something bad is going to happen to you?	4	0.28
7. Do you feel happy most of the time?	2	0.38
8. Do you often feel helpless?	10	0.21
9. Do you prefer to stay at home, rather than going out and doing new things?	N/A	NS
10. Do you feel you have more problems with memory than most?	7	0.23
11. Do you think it is wonderful to be alive now?	N/A	NS
12. Do you feel pretty worthless the way you are now?	N/A	NS
13. Do you feel full of energy?	6	0.19
14. Do you feel that your situation is hopeless?	5	0.22
15. Do you think that most people are better off than you are?	9	0.17

TABLE 5 *10-Item geriatric depression scale (4-item version in bold)*

Are you basically satisfied with your life?	Yes/NO
Have you dropped many of your activities and interests?	YES/No
Do you feel that your life is empty?	YES/No
Are you afraid that something bad is going to happen to you?	YES/No
Do you feel happy most of the time?	Yes/NO
Do you often feel helpless?	YES/No
Do you feel you have more problems with memory than most?	YES/No
Do you feel full of energy?	Yes/NO
Do you feel that your situation is hopeless?	YES/No
Do you think that most people are better off than you are?	YES/No

TABLE 6 *Performance of GDS10, GDS4 and GDS1 against GMS*

	Sensitivity (%)	Specificity (%)	Negative predictive value (%)	Chi-square	P <
GDS15 (cut-off 4/5)	91	72	94	55.1	0.0001
GDS15 (cut-off 5/6)	78	82	87	55.6	0.0001
GDS10 (cut-off 3/4)	89	77	93	61.2	0.0001
GDS10 (cut-off 2/3)	93	63	94	41.2	0.0001
GDS4 (cut-off 1/2)	61	88	80	44.6	0.0001
GDS4 (cut-off 0/1)	93	63	94	41.2	0.0001
GDS1	59	75	77	29.3	0.0001

DISCUSSION

Only 2% of the sample refused to participate, and a further 2% of those interviewed found the GDS15 to be 'not at all' acceptable. It is thus clear that the GDS15 (at least as administered in interview form) is a very feasible screening instrument for elderly people attending their GPs. It also showed very satisfactory internal consistency in this sample. Our further finding that those subjects having difficulty with the questionnaire were very likely to score within the depressed range may itself be a useful clinical pointer.

The primary aim of our study was to assess the feasibility and performance of the GDS15 in elderly GP surgery attenders. The population we interviewed is representative of such attenders and was not intended to reflect the elderly population as a whole. This is borne out by our 'caseness' rate of 34%, which is very similar to that reported in other studies of GP surgery attenders by MacDonald,⁵ who used a brief clinical interview, and by Evans and Katona⁶ who used the longer (30 item) original version of the GDS.¹¹ The consistently higher rate of depression 'caseness' in GP attenders than in epidemiological community surveys such as those by Copeland's group,¹⁷ suggests that elderly depressed patients are more likely than their non-depressed counterparts to visit their GPs frequently and thus be identified early by surgery-based screening. This enhances the potential efficiency of opportunistic screening in picking up depression in elderly attenders of GPs' surgeries.

The similar GDS scores in over and under 75s is consistent with the epidemiological literature.¹⁷ The finding of almost identical 'caseness' rates in men and women is, however, surprising since both GP-based^{5,6} and community¹⁷ studies usually report a higher rate of depression in women. This may simply reflect the relatively low power of a sample of the size of ours to detect a gender difference.

As can be seen from Table 6, we have confirmed in a British primary care sample that the GDS15 has very satisfactory sensitivity (91%) and negative predictive value (95%) against a standardized mental state assessment interview, the GMS, as well as a reasonable specificity of 72%. The present results thus provide some justification for the recent recommendation of the Royal College of General Practitioners¹⁵ that the GDS should be used for routine screening of elderly patients in primary care.

The GDS15 has not, as yet, received wide usage in primary care and is perceived by many GPs as being too long for routine use. The development and demonstrated effectiveness of a very short screening questionnaire could, however, encourage GPs and their ancillary staff to screen a higher proportion of their elderly patients more frequently. With a cut-off of 3/4, our 10 item version has a very similar overall performance to the parent version, with the advantage of omitting some of the redundancy in its content with little loss in terms of internal consistency. The utility of

the GDS10 clearly needs further evaluation, particularly since the saving in time (and therefore acceptability) is small. We have also shown that a single question (do you feel that your life is empty?) identifies nearly two-thirds of the patients diagnosed on detailed psychiatric interview as depressed. We would not, however, recommend routine use of this single question alone since (though much better than nothing) it misses an unacceptably high proportion of depressed patients.

We suggest that the GDS4 deserves particularly careful consideration as a minimal screening procedure for detecting depression in elderly primary care patients. Though its internal consistency was considerably lower than that of the parent scale and it had relatively low specificity, it missed only five of the 46 depressed patients in the validation sample. A two-stage procedure, in which the six extra questions of the GDS10 are given to the subgroup with a score of 1 on the GDS4, may provide the optimal combination of brevity, sensitivity and specificity.

There are clearly limitations to the conclusions that can be drawn from our data on short versions of the GDS. Despite the high correlation between the GDS4 and the longer versions of the scale, there is as yet no evidence that either it or the GDS10 would be satisfactory as measures of severity of depression or to measure change over time. It must also be remembered that our short version GDS data were not obtained directly but were derived from longer interviews in which the GDS15 (derivation sample) and GDS30 (validation sample) were administered. The validation subjects, though drawn from elderly primary care attenders, were not a random sample since a particular attempt was made to interview GDS30 'cases'. Furthermore, GMS interviews were not blind to GDS30 scores and not all patients approached agreed to undergo the lengthy GMS interview. This is reflected in their somewhat higher GDS scores. In demographic characteristics (Table 1), however, they were very similar to the more representative derivation sample.

More fundamental is the issue of whether 'caseness' as detected by such screening procedures represents true depressive illness of a nature and severity warranting and potentially benefitting from treatment interventions. The fact that most of the GDS 'cases' in the validation sample were confirmed by detailed psychiatric interview suggests that they did represent true depression. The relationship with depression as recorded and/or treated by GPs is examined in a companion paper.¹⁸ The community follow-up study by Copeland's group⁴ suggests that only a minority of cases of depression in old age remit spontaneously. We are currently analysing the data from a 1 year follow-up study of the sample reported here to establish whether the same is true for depression as detected by opportunistic screening of elderly primary care attenders.

Despite the limitations of our study, we conclude that extrapolation of our validation results to other elderly primary care patients appears reasonable. The GDS4 in particular merits field testing in a new sample of general practice patients. Such a study could also address the issue of whether effective detection of depression in elderly primary care patients can meet the stringent criteria of a screening programme¹⁹ by enabling more appropriate management and causing measurable improvements in outcome.

ACKNOWLEDGEMENTS

We gratefully acknowledge the help of the doctors, ancillary staff and above all the patients at the practices where the data was collected. We also acknowledge receipt of educational grants from Lilly Industries and Janssen Pharmaceuticals.

REFERENCES

- ¹ Katona CLE. The prognosis of depression in old age. In Arie T (ed.) *Recent advances in psychogeriatrics 2*. Edinburgh: Churchill Livingstone, 1992.
- ² Koenig HG, Shelp F, Veeraindar G, et al. Survival and health care utilisation in elderly medical inpatients with major depression. *J Am Geriatr Soc* 1989; **37**: 599-606.
- ³ Murphy E, Smith R, Lindsay J, Slattery J. Increased mortality rates in late-life depression. *Br J Psychiatry* 1988; **152**: 347-353.
- ⁴ Copeland JR, Davidson IA, Dewey ME, et al. Alzheimer's disease, other dementias, depression and pseudo-dementia: prevalence, incidence and three-year outcome in Liverpool. *Br J Psychiatry* 1992; **161**: 230-239.
- ⁵ MacDonald AJD. Do general practitioners "miss" depression in elderly patients? *Br Med J* 1986; **292**: 1365-1368.
- ⁶ Evans S, Katona CLE. Prevalence of depressive symptoms in elderly primary care attenders. *Dementia* 1993; **4**: 327-333.
- ⁷ Iliffe S, Haines A, Gallivan S, et al. Assessment of elderly people in general practice 1. Social circumstances and mental state. *Br J Gen Pract* 1991; **41**: 9-12.
- ⁸ Livingston G, Hawkins A, Graham N, et al. The Gospel Oak Study: prevalence rates of dementia, depression and activity limitation among elderly residents in inner London. *Psychol Med* 1990; **20**: 137-146.
- ⁹ Freeling P, Rao BM, Paykel ES, et al. Undiagnosed depression in general practice. *Br Med J* 1985; **290**: 1880-1883.
- ¹⁰ Department of Health and the Welsh Office. *General practice in the National Health Service. A new contract*. London: Department of Health and the Welsh Office, 1989.
- ¹¹ Yesavage JA, Brink TL, Rose TL, Lum O. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1983; **17**: 37-49.
- ¹² Katona CLE. Screening for depression: the Geriatric Depression Scale. In Philp I (ed.) *Screening scales for elderly people*, in press.
- ¹³ Copeland JRM, Dewey ME, Griffiths-Jones HM. A computerized psychiatric diagnostic system and case nomenclature for elderly subjects: GMS and AGE-CAT. *Psychol Med* 1986; **16**: 89-99.
- ¹⁴ Sheikh JA, Yesavage JA. Geriatric Depression Scale (GDS): recent findings and development of a shorter version. In Brink TL (ed.) *Clinical gerontology: a guide to assessment and intervention*. New York: Howarth Press, 1986.
- ¹⁵ Williams EI, Wallace P. Health checks for people aged 75 and over. *Br J Gen Pract Occasional Paper No. 59*, 1993.
- ¹⁶ *Statistical Package for the Social Sciences: PC Edition (SPSS/PC+ Version 3.1)*. Chicago: SPSS Inc., 1992.
- ¹⁷ Copeland JRM, Gurland BJ, Dewey ME, et al. Is there more dementia, depression and neurosis in New York? A comparative community study of the elderly in New York and London using the community diagnosis AGE-CAT. *Br J Psychiatry* 1987; **151**: 466-473.
- ¹⁸ Mullan E, D'Ath P, Katona P, Katona C. Screening, detection and management of depression in elderly primary care attenders. II: Detection and fitness for treatment: a case-record study. *Fam Pract* 1994; **11**: 267-270.
- ¹⁹ Holland WW. Screening: reasons to be cautious. *Br Med J* 1993; **306**: 1222-1223.