Yield and costs of direct and stepped screening for depressive symptoms in subjects aged 75 years and over in general practice

Weele, G.M. van der; Waal, M.W.M. de; Hout, W.B. van den; Mast, R.C. van der; Craen, A.J.M. de; Assendelft, W.J.J.; Gussekloo, J.

RESEARCH ARTICLE

International Journal of Geriatric Psychiatry

Yield and costs of direct and stepped screening for depressive symptoms in subjects aged 75 years and over in general practice†

Gerda M van der Weele[1,‡], Margot WM de Waal[1§], Wilbert B van den Hout[2¶], Roos C van der Mast[3‖], Anton JM de Craen[4#], Willem JJ Assendelft[1‖] and Jacobijn Gussekloo[1‖]

[1]Department of Public Health and Primary Care, Leiden University Medical Center, Leiden, the Netherlands
[2]Department of Medical Decision Making, Leiden University Medical Center, Leiden, the Netherlands
[3]Department of Psychiatry, Leiden University Medical Center, Leiden, the Netherlands
[4]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands
Correspondence to: G. M. van der Weele, MD, E-mail: g.m.van_der_weele@lumc.nl

†This paper contains original unpublished work and is not being submitted for publication elsewhere at the same time.

‡Clinical researcher.

§Epidemiologist.

¶Health economist.

‖Professor.

#Senior epidemiologist.

Objective: To examine yield and costs of two screening methods for depressive symptoms in subjects ≥75 years in general practice.

Methods: In 73 general practices of 12.144 registered subjects ≥75 years 10.681 could be invited for screening. In the first 31 practices we invited 3797 subjects for *direct screening* which implied an invitation by letter followed by a home visit to administer the 15-item Geriatric Depression Scale (GDS-15). In the remaining 42 practices 6884 subjects were invited for *stepped screening* which implied that the GDS-15 was sent by post, followed by a home visit only if the self-administered GDS-15-score was ≥ 4 points. Being screen-positive for depressive symptoms was defined as an interviewer-administered GDS-15-score ≥5 points. Screening costs were estimated based on results in this study.

Results: Of all registered subjects 707 (5.8%) were already being treated for depression. The yield of direct screening was higher than of stepped screening (2.6% *versus* 1.9%, $p = 0.009$), with similar yields for subjects aged 75–79 years and for subjects aged ≥80 years. In a standard GP-practice with 160 subjects ≥75 years estimated total screening costs are about twice as high for direct screening than for stepped screening. Estimated costs per screen positive subject are €350 for direct screening and €250 for stepped screening.

Conclusion: Direct screening has a higher yield, but is also more time consuming and more expensive. Whether the extra yield is clinically relevant and worth the extra costs, will depend on the subsequent treatment effect.

Trial registration: www.controlled-trials.com/ISRCTN 71142851

Key words: screening; depressive symptoms; yield; costs; oldest old; primary health care
History: Received 23 November 2009; Accepted 23 February 2010; Published online 21 July 2010 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/gps.2518

Introduction

Depressive symptoms are common in older people and have a negative impact on quality of life (Unutzer *et al*., 2000a), daily functioning (Beekman *et al*., 1997; Chopra *et al*., 2005; Lyness *et al*., 2007), health service utilization (Beekman *et al*., 2002), and mortality (Penninx *et al*., 1999; Cuijpers and Smit, 2002; Stek *et al*., 2005). Furthermore older people with depressive symptoms have an elevated risk to develop major depression (Lyness *et al*., 2006; Schoevers *et al*., 2006; Stek *et al*., 2006). Although depression has been reported to be preventable (van 't Veer-Tazelaar *et al*., 2009), most older people with depressive symptoms do not receive available adequate treatment (Unutzer *et al*., 2000b; Wilson *et al*., 2001; Bijl *et al*., 2004; Van Citters and Bartels, 2004; Skultety and Zeiss, 2006; Cuijpers *et al*., 2007; Steinman *et al*., 2007). Under-treatment of depressive symptoms in older people may partly be due to poor recognition because of somatic co-morbidity and because older people, their family and their general practitioners (GP) easily regard depressive symptoms as normal in the process of aging (Crawford *et al*., 1998; Volkers *et al*., 2004; Murray *et al*., 2006). To enhance case detection, screening for depressive symptoms is recommended, if treatment is available for screen-positive subjects (Valenstein *et al*., 2001; Pignone *et al*., 2002). However, the advisability of screening programs is still debated because evidence on the effectiveness of screening programs for depression is inconclusive. Some large trials showed positive results with respect to depressive symptoms among intervention patients who were (partly) detected by screening (Unutzer *et al*., 2002; Bruce *et al*., 2004), whereas other trials found no positive effects (Callahan *et al*., 1994; Whooley *et al*., 2000).

In research settings, different screening approaches have been used such as depression screening integrated in comprehensive health checks (Arthur *et al*., 2002; Osborn *et al*., 2002; van 't Veer-Tazelaar *et al*., 2008) *versus* screening for depression only (Licht-Strunk *et al*., 2005), screening all subjects registered in a general practice (van 't Veer-Tazelaar *et al*., 2008) *versus* screening attendees only (Licht-Strunk *et al*., 2005; Weyerer *et al*., 2008), and interviewing all responders (Arthur *et al*., 2002; Osborn *et al*., 2002; Luppa *et al*., 2008) *versus* two-stepped screening in which subjects are interviewed only after screening positive on a self-administered questionnaire (Licht-Strunk *et al*., 2005). Since in the Netherlands everyone has a GP, systematic screening of all community dwelling older people is suitable.

In order to enhance implementation in daily practice, evidence based information is needed about both screening and treatment of screen-positive subjects. Especially for the oldest old there is not enough information about the yield of screening in relation to the costs. Therefore, we studied the differences in yield and costs of two methods of screening for the broad spectrum of depressive symptoms in all subjects aged 75 years and over who were registered in general practice, thereby also including those who (almost) never visit their GP. Firstly, we chose a screening method that theoretically would ensure the highest yield by visiting all participants at their own home. Secondly, we chose a less labor intensive and time consuming method, in which not all participants had to be visited. We hypothesized that the more intensive, costly screening method would result in a higher yield, especially among the oldest old.

Consequently, our research questions were 'What are the differences in yield and costs of two methods of screening for untreated depressive symptoms in subjects aged 75 years and over in general practice; and do costs and yield differ between subjects aged 75–79 years and subjects $\geq$ 80 years?'

Methods

Study population and procedures

This screening study is part of the PROMODE study (PROactive Management Of Depression in the Elderly) and was followed by a pragmatic, cluster-randomised, controlled trial. The Medical Ethical Committee of the Leiden University Medical Center approved of the study.

We invited all GP practices with at least 80 registered subjects aged $\geq$ 75 years in the Leiden region, the Netherlands ($n = 113$): 73 practices (65%) consented to participate and 40 (35%) not. The study population consisted of all subjects aged $\geq$ 75 years, registered in the 73 participating practices. Exclusion criteria were current treatment for depression (psychological and/or use of antidepressants), a clinical diagnosis of dementia or a Mini-Mental State Examination-score (MMSE) $<$ 19 points (Folstein *et al*., 1975), loss of partner or child in the preceding 3 months, terminal illness with a life expectancy $<$3 months and not speaking Dutch.

From April 2007 to July 2008, subjects were invited to participate (see flow chart, figure 1). For direct screening, all subjects in the first 31 consecutive

practices were invited to participate by sending them a study information leaflet and a response card for agreement to be contacted for a home visit. After 2–3 weeks non-responders were phoned to ask for participation. Trained research nurses visited all subjects who agreed for an interview. For stepped screening, in the remaining 42 consecutive practices we sent the 15-item Geriatric Depression Scale (GDS-15) by post, with a postal reminder after 2–3 weeks. Subjects who scored ≥4 points on the self-administered GDS-15 were visited for an interview by a research nurse. In order to get an indication of false negatives, all subjects with a self-administered GDS-15-score <4 of the first practice receiving stepped screening were also interviewed. Of 72 subjects only 1 had an interviewer-administered GDS-15-score ≥5 (negative predictive value = 0.98). In all interviews the MMSE and GDS-15 were administered. In screen-positives (interviewer-administered GDS-15 scores ≥5 points), we administered also the Montgomery Åsberg Depression Rating Scale (MADRS) for assessment of severity of depression and the Mini-International Neuropsychiatric Interview (MINI) for assessment of DSM-IV diagnoses of depression.

Measurements

The GDS-15, used as screening instrument, has been developed to screen for depression in older populations (Sheikh and Yesavage, 1986). In a recent meta-analysis (based on 69 studies) of the diagnostic accuracy, clinical utility and added value of the GDS-15 in primary care, the GDS-15 was rated 'good' for screening (Mitchell *et al*., 2009). It ranges from 0 to 15 points, with higher scores indicating more depressive symptoms. We considered an interviewer-administered GDS-score ≥ 5 points as screen-positive, because we chose to include the broad depressive spectrum, ranging from symptoms not meeting formal DSM-IV diagnostic criteria for depression to major depression. The cut-off score of ≥5 points that we used to define being screen-positive provides good sensitivity and specificity in primary care settings (D'Ath *et al*., 1994; De Craen *et al*., 2003).

Although diagnosing depressive disorders was not the focus of this study, we were interested whether both methods yielded screen-positive subjects with similar presence of a depressive disorder according to formal DSM-IV-criteria and similar severity of depression. To assess whether screen-positive subjects had a depressive disorder according to formal DSM-IV-criteria, including minor and major depression and dysthymia, we used the MINI (Sheehan *et al*., 1998). To assess the severity of depression among screen-positive subjects we used the MADRS. The MADRS is an interviewer-rated scale that can be administered by trained nurses. The MADRS consists of 10 items ranging from 0–6 points, resulting in possible total scores of 0–60 points. Higher scores indicate more severe depression.

To assess possible differences in presence of depressive symptoms between participants and non-participants in the screening program, GPs were asked to give their clinical judgment for each invited subject about the presence of depressive symptoms, before inclusion started. The possible answers were 'present, possibly present, absent, or unknown'.

Training of research nurses

Training of nurses occurred on a regular basis. The nurses were instructed how to use the different questionnaires in a uniform way, with special attention to the GDS and MADRS. For the GDS, which has a yes/no-answering format with very strict standard questions, training concentrated mainly on the use of the exact phrasing of questions. With respect to the MADRS all interviews were videotaped. These tapes were used for both training and inter-rater agreement. If no initial agreement existed discussion followed to reach consensus.

Statistical analysis

The yield and costs of the two different screening methods were compared within subjects aged 75–79 years and subjects aged ≥80 years and for both sexes, separately. Data are presented in absolute numbers with percentages and in medians with interquartile ranges (IQR). Categorical data were analyzed using χ^2 test and continuous data using Mann–Whitney *U*-test. Statistical analyses were carried out using the SPSS 16.0 for Windows.

Economic evaluation

We calculated expected costs if screening would be implemented in a standard Dutch general practice with 160 subjects ≥ 75 years (80 subjects aged 75–79 years and 80 subjects aged ≥80 years) given the exclusion rate, participation rate and screening yield found in our study, with the additional assumption that 20% of interviews needs to be done at home in subjects

75–79 years and 40% in subjects ≥80 years. Time investment for the different procedures was estimated on experience in our study. We used the 2009 Dutch tariffs for the staff members in general practice who would be involved, including administrative procedures by the practice assistant (PA) and administration of GDS-15 by a practice nurse (PN) (www.NVDA.nl, 2009; www.Nza.nl, 2009).

Results

Study population

Of all subjects aged 75 years and over who were registered in the participating practices, 1463 subjects (12%) were excluded by their GPs (Figure 1), mostly because of current depression treatment ($n = 707$) or a clinical diagnosis of dementia or MMSE < 19 points ($n = 631$). For direct screening, 3797 subjects were invited, of whom 166 (4%) were excluded after invitation (49 before the interview took place + 117 during baseline interviews). Of the remaining subjects 54% (1965/3631) participated. For stepped screening, 6884 subjects were invited, of whom 140 (2%) were excluded after invitation (78 before the interview took place + 62 during baseline interviews). Of the remaining subjects 50% (3365/6744) participated.

The percentage of subjects aged ≥80 years in direct screening was somewhat higher than in stepped screening (56% *versus* 52%, $p < 0.001$) (Table 1). In direct screening the prevalence of exclusion criteria was higher than in stepped screening (17% *versus* 13%, $p < 0.001$), mainly due to a higher percentage of 'other exclusion criteria', such as not speaking Dutch. Populations in both screening methods did not differ regarding to sex, exclusion for current depression treatment and clinical diagnosis of dementia and GP's

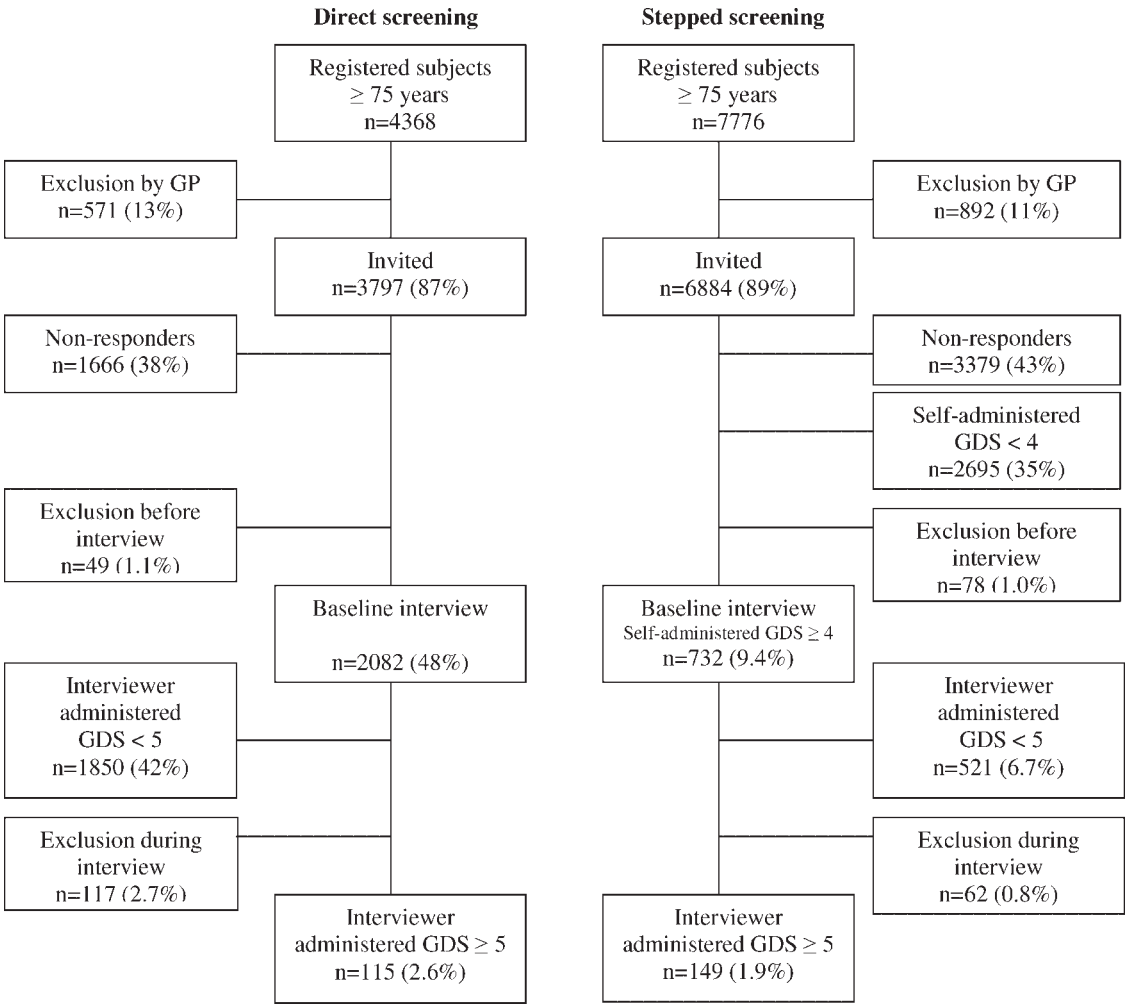

Figure 1 Flow chart screening (numbers of subjects and % of registered subjects).

Table 1 Characteristics of registered subjects per screening method in total study population ($n = 12\,144$)

	Direct screening ($n = 4368$)		Stepped screening ($n = 7776$)		p-value
	n	%	n	%	
Age					
75–79 years	1937	(44)	3711	(48)	< 0.001
≥ 80 years	2431	(56)	4065	(52)	
Sex					
Female	2729	(63)	4780	(62)	0.27
Male	1639	(38)	2996	(39)	
Exclusion criteria present	737	(17)	1032	(13)	< 0.001
Current depression treatment	282	(6.5)	425	(5.5)	0.32
Clinical diagnosis of dementia /MMSE < 19	235	(5.4)	396	(5.1)	
GP-judgment '(possibly) depressed' without exclusion	785	(18)	1195	(17)	0.33

clinical judgment of depression, also if data were analyzed for both age groups separately (data not shown).

Yield of screen-positives

In the total sample of registered subjects, the yield of direct screening was higher than of stepped screening (2.6% (115/4368) *versus* 1.9% (149/7776), $p = 0.009$), without difference between age groups (Table 2). Analyses at the GP-level also gave higher yields of direct screening compared to stepped screening (2.7% (range 0–5%) *versus* 1.8% (range 0–4%), $p = 0.003$). The yield of direct and stepped screening was similar for men (1.9% *versus* 1.6%, $p = 0.51$). For women however, the yield of direct screening was higher than the yield of stepped screening (3.1% *versus* 2.1%, $p = 0.009$). This was seen in both age groups, but was most clear in the age group ≥ 80 years (data not shown).

Table 2 Yield of screen-positive subjects (GDS-15 ≥ 5) per screening method in total study population ($n = 12\,144$)

	Direct screening		Stepped screening		p-value
	Yield[a]	%	Yield[a]	%	
Age					
75–79 years	51/1937	(2.6)	69/3711	(1.9)	0.056
≥ 80 years	64/2431	(2.6)	80/4065	(2.0)	0.078
Sex					
Male	30/1639	(1.9)	47/2996	(1.6)	0.51
Female	85/2729	(3.1)	102/4780	(2.1)	0.009
Total	115/4368	(2.6)	149/7776	(1.9)	0.009

GDS-15 = Geriatric Depression Scale 15-items version.
[a]Yield: screen positive subjects per (sub)group.

Comparison between participants and non-participants

In Table 3, participants and non-participants are compared according to sex and GP clinical judgment on the presence of depression for both age groups and screening methods separately. In direct screening, for both age groups there was no difference in sex distribution. In stepped screening however, in both age groups participants compared to non-participants were less likely to be women. According to GP clinical judgment of depression, in direct screening participants compared to non-participants were less likely to be judged as depressed or possibly depressed. In stepped screening this was also seen in the age group ≥80 years, but not in the age group 75–79 years (Table 3).

Characteristics of screen-positive subjects

To explore if screen-positive subjects in both screening methods had similar characteristics, we could evaluate data of 239 subjects. Of all 264 screen-positive subjects 239 agreed to additional assessment of depression diagnosis and severity; in direct screening 13 subjects could not be further assessed (2 dropped out, 11 refused) and in stepped screening 12 subjects (2 dropped out, 10 refused). Screen-positive subjects did not differ between screening methods with respect to demographic characteristics, including age, sex and living alone. Severity of depressive symptoms and presence of DSM-IV depression diagnoses were also similar. Median MADRS-score was 13 points in both groups. In direct screening a DSM-IV diagnosis was present in 41% of screen-positive subjects, in stepped screening this was 46% ($p = 0.46$).

Table 3 Participants and non-participants of both screening methods ($n = 10375$[a])

	Direct screening					Stepped screening				
	Participants		Non-participants		*p*-value	Participants		Non-participants		*p*-value
	n	%	*n*	%		*n*	%	*n*	%	
75–79 years										
Sex										
Male	403	(44)	331	(43)	0.77	888	(50)	593	(38)	<0.001
Female	512	(56)	433	(57)		900	(50)	952	(62)	
GP clinical judgment										
(Possibly) depressed	164	(18)	181	(24)	0.004	319	(18)	288	(19)	0.55
Not depressed[b]	751	(82)	583	(76)		1469	(82)	1257	(81)	
Total	915		764			1788		1545		
≥80 years										
Sex										
Male	390	(37)	310	(34)	0.20	626	(40)	564	(31)	<0.001
Female	660	(63)	592	(66)		951	(60)	1270	(69)	
GP clinical judgment										
(Possibly) depressed	200	(19)	240	(27)	<0.001	254	(16)	358	(20)	0.010
Not depressed[b]	850	(81)	662	(73)		1323	(84)	1476	(81)	
Total	1050		902			1577		1834		

[a]all registered subjects (n=12144) minus excluded subjects (n=1769).
[b]not depressed, including judgment 'unknown' or missing.

Estimated costs and cost effectiveness of both screening methods

Table 4 shows the estimated costs of both screening methods if screening would be implemented in a standard GP-practice with 80 subjects 75–79 years and 80 subjects ≥80 years. Compared to stepped screening, per standard practice total costs of direct screening would be about twice as high (75–79 years €745 *versus* €374, ≥80 years €764 *versus* €349, total €1509 *versus* €723), since the number of contacts for interviews was much higher in direct screening. Estimated costs per screen-positive subject were also higher for direct screening than for stepped screening, about €350 *versus* €250 (75–79 years €346 *versus* €249, ≥80 years €349 *versus* €224). With direct screening 0.7% additional screen-positives were found for €569 per additional screen-positive in the age group 75–79 years and €659 in the age group ≥80 years.

Discussion

In our study the yield of screening for untreated depressive symptoms among all registered subjects aged 75 years and over in general practice was 2.2%, with a somewhat higher yield in direct screening than in stepped screening. With direct screening total screening costs and costs per registered subject are about twice as high as with stepped screening.

In this paper the yield of screening was defined as the percentage of registered subjects that screened positive. In general, interpretation of the magnitude of differences in yield cannot be separated from whether subjects will subsequently agree to undergo treatment and whether treatment will be beneficial. For instance, when all detected subjects will benefit from treatment, even a 1% difference in yield may be regarded as clinically relevant. When only 10% of detected subjects will benefit from treatment, a 1% difference in yield is not likely to be clinically relevant. The yield of both screening and treatment together will determine the clinical relevance of the screening method. Our paper gives information on the crucial first step in this process. Regarding treatment effects among screen-positive depressive subjects, both positive and negative results have been reported. Cuijpers *et al*. recently published a review of trials that examined the effects of psychological treatment; studies with extra elements (like managed care interventions or disease management programs) were excluded. They concluded that psychological treatment of depression has little effect when patients are recruited through systematic screening (as opposed to patients who are referred for

Table 4 Estimated costs of direct and stepped screening per age group in a standard GP-practice with 80 subjects 75-79 years and 80 subjects ≥ 80 years

	Staff time[a]	75–79 years				≥ 80 years			
		Direct screening		Stepped screening		Direct screening		Stepped screening	
	Minutes and type of staff	*n*	Estimated costs in €	*n*	Estimated costs in €	*n*	Estimated costs in €	*n*	Estimated costs in €
a. Basic costs[b]		80	249	80	285	80	235	80	268
Number of subjects invited		71.2		71.2		66.4		66.4	
b. Scoring self-administered GDS	1 PA	—	—	38.4	17	—	—	30.5	14
c. Making interview appointment	3 PA	38.4	52	4.2	6	35.9	48	3.4	5
d. Interview[c]									
-in GP-practice	15 PN	30.8	277	3.4	30	21.5	194	2.0	18
-at home	30 PN	7.7	138	0.8	15	14.3	258	1.3	24
e. Evaluation screen-positive result	5 PN + GP	2.2	16	1.5	11	2.2	16	1.6	12
f. Exchange results with patient	10 PN	2.2	13	1.5	9	2.2	13	1.6	9
Total costs per practice			745		374		764		349
Costs per registered subject			9		5		10		5
Costs per screen positive			346		249		349		224
Costs per additional screen positive (direct *versus* stepped screening)			569				659		

GDS = Geriatric Depression Scale, 15-item-version.
Estimated costs = staff time in minutes times amount in € per minute staff time times number of subjects left (*n*).
[a]PA = Practice Assistant €0.45 per minute; PN = Practice Nurse €0.60 per minute; GP = General Practitioner €0.90 per minute.
[b]Basic costs, based on observed exclusion and response rates, include: material costs and staff time for extraction electronic patient list by PA, check for exclusion criteria by GP, invitation by post and reminder (per telephone or postal) by PA
[c]assumption: in group 75–79 years 20% of screening interviews have to be performed at home and in group ≥80 years 40%.

treatment by their GP) (Cuijpers *et al*., 2009). However, encouraging results were found in the PROSPECT trial, in which care management that was offered to screen-positive subjects ≥60 years reduced suicidal ideation (Bruce *et al*., 2004), and in a recent Dutch trial, in which stepped care offered to screen-positive subjects ≥75 years reduced the risk to develop a depressive or anxiety disorder (van 't Veer-Tazelaar *et al*., 2009).

Our overall yield of 2% is lower than the 7–36% reported in earlier screening studies (Koenig and Blazer, 1992). Two very similar studies, that also screened subjects aged ≥75 years in general practice for depressive symptoms defined as a GDS-15 score ≥5 points, showed yields of 8% (Arthur *et al*., 2002) and 13% (Osborn *et al*., 2002). However, in contrast to these studies, we excluded 6% of all registered subjects because they already received depression treatment. Taking this into account, our yield is in line with earlier findings.

Participation in direct and stepped screening was very similar (54% and 50%, respectively). Apparently a more labor-intensive approach does not guarantee higher response rates. Beforehand we had no hypothesis on gender differences. We found that participation among women was higher in direct screening than in stepped screening. Among men we found no difference in yield of screen-positives between the two screening methods. Among women the yield of screen-positives was 1% higher in direct screening than in stepped screening, but again the clinical relevance of this finding can be debated.

Costs of screening differed with total number and participation rate of subjects, the payment of the screening staff member and the number of subjects to be visited at home. Costs differ mostly with the number of subjects to be interviewed, which is related to the screening method. Compared to direct screening, pre-screening by self-administered GDS-15, possibly followed by a phone call to non-responders for assessment of GDS, will reduce costs. Further reduction in costs might be possible by integrating screening in routine clinical practice, since most older people have several contacts every year with their GP. However, implementation of pro-active case finding during regular contacts often proved to be difficult in Dutch general practice (www.nationaalkompas.nl, 2009).

A major strength of this study is the comparison of yield and costs of two different screening methods for untreated depressive symptoms in a large number of subjects aged 75 years and over. In addition, this study gives insight in the (un)willingness of older people to participate in screening for depressive symptoms depending on method of invitation, age, and sex. Our findings can help in making decisions how screening for depressive symptoms in older age groups should be done. Because there is increasing evidence that treatment of depressive symptoms in late life can have important positive effects, such as preventing the onset of major depression (van 't Veer-Tazelaar *et al*., 2009), we screened for depressive symptoms rather than depressive disorders. A next step to investigate should be whether people with depressive symptoms who are detected by screening gain as much from (preventive) interventions as people who present themselves with symptoms.

It could be seen as a limitation that practices were not randomized. Over a period of 14 months all GP practices, with at least 80 registered subjects aged $\geq$75 years, in the Leiden region were invited to participate. No other selection criteria were used. In the first 31 practices that were included direct screening was performed and in the next 42 practices stepped screening. During the recruitment period rates and reasons to refuse participation did not change; the reasons that were mentioned were lack of time and being engaged in practice reorganization or in other research projects. Because of the unselective recruitment combined with the high number of participating practices, we had no reason to believe that practice characteristics differ between both screening methods regarding to prevalence of depressive symptoms in their population and the patients' willingness to be screened. Furthermore, our findings that the populations in both screening methods did not differ regarding current depression treatment and GP's clinical judgment of depression did not raise suspicion that GPs in both methods differed in recognizing and treating depression at high age.

In our study the response rates were 50–54%. Low response rates combined with high exclusion rates could compromise the validity and generalizability of results in a prevalence study. This is, however, not the case for our pragmatic study, in which we assessed the differences in yield and costs as outcome measures for the two different screening strategies. Similarly, differences between participants and non-participants do not undermine our results, but show characteristics of the particular screening strategies. Because selective non-response of depressed subjects could undermine cost-effective screening we were especially interested whether non-responders were more frequently judged as depressed by their GP compared to responders. Mainly in direct screening, we found that subjects who were judged by their GP as depressed or possibly depressed participated less frequently than subjects who were not depressed. This suggests that a more intensive approach does not necessarily attract those subjects the GPs most worry about. Despite this, the yield of direct screening was higher. We hypothesize that non-responders could be the ones less likely to accept an unsolicited treatment offer.

In stepped screening we introduced self-administration of the GDS, but being screen-positive was still based on the interviewer-administered GDS. O'Neill *et al*. found that self-administered GDS-scores are generally higher than interviewer-rated scores of the same person (O'Neill *et al*., 1992). Therefore, not visiting subjects with a self-administered GDS-score <4 points seems acceptable. Even if some of the GDS-scores would be incorrect, this still would not invalidate our pragmatic comparison of the two screening strategies. Self-administration of the GDS also introduces the possibility that a proxy fills out the form. We expect that subjects with visual impairment or illiteracy would have needed help from a proxy to be able to participate in screening. Problems could arise when a proxy fills in the GDS-form without consulting the person who was intended to be screened, but filling out depression questionnaires by a proxy generally results in higher scores compared to self-administered scores (Brown and Schinka, 2005; Nitcher *et al*., 2009).

In conclusion, our study shows that direct screening of all registered subjects aged 75 years and over in a GP-practice for untreated depressive symptoms has a higher yield than stepped screening. However, total screening costs per practice of direct screening are twice as high as total costs of stepped screening. Whether the extra yield is clinically relevant and worth the extra costs will depend on the subsequent treatment effect.

Conflict of interest

None declared.

Author contributions

Study concept and design was done by de Waal, van den Hout, van der Mast, Assendelft, and Gussekloo. Acquisition of data was carried out by van der Weele,

Key Points

- In this study the yield of screening for untreated depressive symptoms among all registered subjects aged 75 years and over in general practice was 2.2%.
- Direct screening has a higher yield than stepped screening, but costs are about twice as high for direct screening.
- Whether the extra yield is clinically relevant and worth the extra costs, will depend on the subsequent treatment effect.

de Waal, van der Mast and Gussekloo. Analysis and interpretation of data was carried out by van der Weele, de Waal, van den Hout, van der Mast, de Craen, Assendelft, and Gussekloo. Manuscript drafting was done by van der Weele, de Waal, van der Mast, and Gussekloo. Critical revision of the paper for important intellectual content was done by van der Weele, de Waal, van den Hout, van der Mast, de Craen, Assendelft and Gussekloo, whereas final approval of the version to be published was given by van der Weele, de Waal, van den Hout, van der Mast, de Craen, Assendelft and Gussekloo.

Acknowledgements

The study was funded by a grant from the Netherlands Organisation for Health Research and Development (ZonMw; grant number: 945-07-502) and a grant from Rivierduinen, Center for mental health care.

REFERENCE

Arthur AJ, Jagger C, Lindesay J, Matthews RJ. 2002. Evaluating a mental health assessment for older people with depressive symptoms in general practice: a randomised controlled trial. *Br J Gen Pract* **52**(476): 202–207.

Beekman AT, Deeg DJ, Braam AW, Smit JH, van Tilburg W. 1997. Consequences of major and minor depression in later life: a study of disability, well-being and service utilization. *Psychol Med* **27**(6): 1397–1409.

Beekman AT, Penninx BW, Deeg DJ, *et al*. 2002. The impact of depression on the well-being, disability and use of services in older adults: a longitudinal perspective. *Acta Psychiatr Scand* **105**(1): 20–27.

Bijl D, van Marwijk HW, de Haan M, van Tilburg W, Beekman AJ. 2004. Effectiveness of disease management programmes for recognition, diagnosis and treatment of depression in primary care. *Eur J Gen Pract* **10**(1): 6–12.

Brown LM, Schinka JA. 2005. Development and initial validation of a 15-item informant version of the Geriatric Depression Scale. *Int J Geriatr Psychiatry* **20**(10): 911–918.

Bruce ML, Ten Have TR, Reynolds CF III, *et al*. 2004. Reducing suicidal ideation and depressive symptoms in depressed older primary care patients: a randomized controlled trial. *JAMA* **291**(9): 1081–1091.

Callahan CM, Hendrie HC, Dittus RS, *et al*. 1994. Improving treatment of late life depression in primary care: a randomized clinical trial. *J Am Geriatr Soc* **42**(8): 839–846.

Chopra MP, Zubritsky C, Knott K, *et al*. 2005. Importance of subsyndromal symptoms of depression in elderly patients. *Am J Geriatr Psychiatry* **13**(7): 597–606.

Crawford MJ, Prince M, Menezes P, Mann AH. 1998. The recognition and treatment of depression in older people in primary care. *Int J Geriatr Psychiatry* **13**(3): 172–176.

Cuijpers P, Smit F. 2002. Excess mortality in depression: a meta-analysis of community studies. *J Affect Disord* **72**(3): 227–236.

Cuijpers P, Smit F, van Straten A. 2007. Psychological treatments of subthreshold depression: a meta-analytic review. *Acta Psychiatr Scand* **115**(6): 434–441.

Cuijpers P, van Straten A, van Schaik A, Andersson G. 2009. Psychological treatment of depression in primary care: a meta-analysis. *Br J Gen Pract* **59**(559): e51–e60.

D'Ath P, Katona P, Mullan E, Evans S, Katona C. 1994. 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* **11**(3): 260–266.

De Craen AJ, Heeren TJ, Gussekloo J. 2003. Accuracy of the 15-item geriatric depression scale (GDS-15) in a community sample of the oldest old. *Int J Geriatr Psychiatry* **18**(1): 63–66.

Folstein MF, Folstein SE, McHugh PR. 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* **12**(3): 189–198.

Koenig HG, Blazer DG. 1992. Epidemiology of geriatric affective disorders. *Clin Geriatr Med* **8**(2): 235–251.

Licht-Strunk E, van der Kooij KG, van Schaik DJ, *et al*. 2005. Prevalence of depression in older patients consulting their general practitioner in The Netherlands. *Int J Geriatr Psychiatry* **20**(11): 1013–1019.

Luppa M, Heinrich S, Matschinger H, *et al*. 2008. Direct costs associated with mild cognitive impairment in primary care. *Int J Geriatr Psychiatry* **23**(9): 963–971.

Lyness JM, Heo M, Datto CJ, *et al*. 2006. Outcomes of minor and subsyndromal depression among elderly patients in primary care settings. *Ann Intern Med* **144**(7): 496–504.

Lyness JM, Kim J, Tang W, *et al*. 2007. The clinical significance of subsyndromal depression in older primary care patients. *Am J Geriatr Psychiatry* **15**(3): 214–223.

Mitchell AJ, Bird V, Rizzo M, Meader N. 2009. Diagnostic validity and added value of the geriatric depression scale for depression in primary care: a meta-analysis of GDS(30) and GDS(15). *J Affect Disord* DOI:10.1016/j.jad.2009.08.019

Murray J, Banerjee S, Byng R, *et al*. 2006. Primary care professionals' perceptions of depression in older people: a qualitative study. *Soc Sci Med* **63**(5): 1363–1373.

Nitcher RL, Burke WJ, Roccaforte WH, Wengel SP. 2009. A collateral source version of the geriatric depression rating scale. *Am J Geriatr Psychiatry* **1**(2): 143–152.

O'Neill D, Rice I, Blake P, Walsh B, Coakley D. 1992. The geriatric depression scale: rater-administered or self-administered? *Int J Geriatr Psychiatry* **7**(7): 511–515.

Osborn DP, Fletcher AE, Smeeth L, *et al*. 2002. Geriatric depression scale scores in a representative sample of 14 545 people aged 75 and over in the United Kingdom: results from the MRC trial of assessment and management of older people in the community. *Int J Geriatr Psychiatry* **17**(4): 375–382.

Penninx BW, Geerlings SW, Deeg DJ, *et al*. 1999. Minor and major depression and the risk of death in older persons. *Arch Gen Psychiatry* **56**(10): 889–895.

Pignone MP, Gaynes BN, Rushton JL, *et al*. 2002. Screening for depression in adults: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med* **136**(10): 765–776.

Schoevers RA, Smit F, Deeg DJ, *et al*. 2006. Prevention of late-life depression in primary care: do we know where to begin? *Am J Psychiatry* **163**(9): 1611–1621.

Sheehan DV, Lecrubier Y, Sheehan KH, *et al*. 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* **59** (Suppl 20): 22–33.

Sheikh JI, Yesavage JA. 1986. Geriatric depression scale (GDS): recent evidence and development of a shorter version. In *Clinical Gerontology; A Guide to Assessment and Intervention*, Brink LT (ed.). Howarth Press: New York.

Skultety KM, Zeiss A. 2006. The treatment of depression in older adults in the primary care setting: an evidence-based review. *Health Psychol* **25**(6): 665–674.

Steinman LE, Frederick JT, Prohaska T, *et al*. 2007. Recommendations for treating depression in community-based older adults. *Am J Prev Med* **33**(3): 175–181.

Stek ML, Vinkers DJ, Gussekloo J, *et al*. 2005. Is depression in old age fatal only when people feel lonely? *Am J Psychiatry* **162**(1): 178–180.

Stek ML, Vinkers DJ, Gussekloo J, *et al*. 2006. Natural history of depression in the oldest old: population-based prospective study. *Br J Psychiatry* **188**: 65–69.

Unutzer J, Katon W, Callahan CM, *et al*. 2002. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA* **288**(22): 2836–2845.

Unutzer J, Patrick DL, Diehr P, *et al*. 2000a. Quality adjusted life years in older adults with depressive symptoms and chronic medical disorders. *Int Psychogeriatr* **12**(1): 15–33.

Unutzer J, Simon G, Belin TR, *et al*. 2000b. Care for depression in HMO patients aged 65 and older. *J Am Geriatr Soc* **48**(8): 871–878.

Valenstein M, Vijan S, Zeber JE, Boehm K, Buttar A. 2001. The cost-utility of screening for depression in primary care. *Ann Intern Med* **134**(5): 345–360.
Van Citters AD, Bartels SJ. 2004. A systematic review of the effectiveness of community-based mental health outreach services for older adults. *Psychiatr Serv* **55**(11): 1237–1249.
Van 't Veer-Tazelaar PJ, van Marwijk HW, Jansen AP, *et al.* 2008. Depression in old age (75+), the PIKO study. *J Affect Disord* **106**(3): 295–299.
Van 't Veer-Tazelaar PJ, van Marwijk HW, van Oppen P, *et al.* 2009. Stepped-care prevention of anxiety and depression in late life: a randomized controlled trial. *Arch Gen Psychiatry* **66**(3): 297–304.
Volkers AC, Nuyen J, Verhaak PF, Schellevis FG. 2004. The problem of diagnosing major depression in elderly primary care patients. *J Affect Disord* **82**(2): 259–263.
Weyerer S, Eifflaender-Gorfer S, Kohler L, *et al.* 2008. Prevalence and risk factors for depression in non-demented primary care attenders aged 75 years and older. *J Affect Disord* **111**(2–3): 153–163.
Whooley MA, Stone B, Soghikian K. 2000. Randomized trial of case-finding for depression in elderly primary care patients. *J Gen Intern Med* **15**(5): 293–300.
Wilson K, Mottram P, Sivanranthan A, Nightingale A. 2001. Antidepressant versus placebo for depressed elderly. *Cochrane Database Syst Rev* (2): CD000561.
www.nationaalkompas.nl (consulted 15-2-2010). Volksgezondheid Toekomst Verkenning, Nationaal Kompas Volksgezondheid. Bilthoven: RIVM. version 3.20, December 2009.
www.NVDA.nl (consulted 15-12009). 2009.
www.Nza.nl (consulted 15-1-2009). 2009. Tarievenlijst huisartsenzorg, Nederlandse Zorgautoriteit.

