Chapter 2

A high response is not essential to prevent selection bias: results from the Leiden 85-plus Study

A. Bootsma-van der Wiel (1); E. van Exel (1); A.J.M. de Craen (1,2); J.Gussekloo (1),

A.M. Lagaay (1), D.L. Knook (1), R.G.J. Westendorp (1)

Gerontology and Geriatrics, department of General Internal Medicine (1) and Clinical Epidemiology (2), Leiden University Medical Center, the Netherlands.

In press J Clin Epid

ABSTRACT

Background We tested the hypothesis that an additional effort to increase the response rate would diminish selection bias in a community-based cohort study.

Methods In the Leiden 85-plus Study, all subjects of the town of Leiden who had reached their 85th birthday were informed of the study by mail and then asked to participate by telephone. In an additional recruitment stage, those subjects who did not participate directly were visited and personally asked to participate. When these subjects refused, some non-response questions were asked. In this way we collected data on the whole source population.

Results Of 691 eligible elderly subjects, 511 subjects (74%) participated directly. Of those who did not participate directly, 88 subjects participated after the additional effort. The response rate increased from 74% to 87%. Compared to the 511 subjects who directly participated, the 88 subjects who entered the study after the additional effort had poorer health and lower survival. The subjects who refused were more healthy and had poorer mood. The direct sample did not differ from the source population with respect to socio-demographics, health, and mortality.

Conclusion We showed that given a moderately high direct response the additional effort was effective in increasing the response rate, but was also selective and was not necessary to prevent selection bias.

Introduction

A highly representative sample of participants is no longer considered essential for generalisability in etiological studies that report risk estimates rather than prevalence estimates[1-3]. Even a minimum of 80 percent response in follow-up studies is debatable[4]. Generalisability depends on the ability to abstract universal scientific hypotheses or theories from a set of observations and not only from the statistical framework of these observations[5,6]. However, many studies in the elderly have a public health goal in addition to more scientific etiological goals. In such community surveys, generating estimates that can be extrapolated to the general population, representativeness is still very important. Furthermore, it is essential to include frail elderly subjects in a study to investigate the determinants of and causal relations with chronic conditions. Refusal to participate due to ill health would surely invalidate results on the impact of chronic conditions in an elderly population[7] .

A high response rate increases the validity of community-based studies, since a low response rate might lead to selection bias[8]. The success of the response depends to a great extent on the way eligible subjects are approached. A high response can be achieved by interviewing and examining elderly subjects in their homes, since frail elderly subjects are less inclined to visit a study site[9]. Other effective strategies to optimise response rates are notification in advance by mail, involvement of expert researchers, and the prospect of a small gift[10]. Another possibility to increase response is to approach eligible subjects who initially declined or did not respond at all[11,12]. Using these strategies surveys among the elderly have been conducted resulting in response rates between 60 and 90 percent[13-24]. Differences in characteristics and associations between the sample of participants and the source population, however, frequently remain unknown.

In the Leiden 85-plus Study, a research nurse visited all subjects who did not participate directly after the first approach by telephone. Through this additional effort more subjects were drawn into the study. Moreover, the nurse asked a few questions to those who refused to participate to get an impression of their health and well-being. In this way we collected data from the whole source population. This provided an excellent opportunity to test the hypothesis that the additional effort to increase the response rate had diminished selection bias.

Methods

The Leiden 85-plus Study is a series of gerontological surveys of the population of the oldest old living in the town of Leiden, the Netherlands. The first survey started in 1986. The present survey is a community-based follow-up study in a delineated cohort of 85-year-olds. Special topics within the Leiden 85-plus Study are atherosclerosis, cognitive function, chronic diseases, disabilities, and well-being.

Study design

Between September 1997 and September 1999, all members of the 1912 to 1914-birth cohort (n = 705) were eligible to participate in the study. Subjects of the source population were informed about

the study by mail in the week after their 85th birthday. Within a month a physician or research nurse contacted them by phone to request their participation. If subjects agreed to participate, they were visited at their place of residence, oral and written information about the study was provided, and oral informed consent was obtained. When participants were severely cognitively impaired, informed consent was obtained from a responsible person.
When subjects hesitated or declined during the first telephone contact or when they could not be reached by phone, they entered the additional recruitment stage. In this stage the research nurse approached the subjects at their place of residence. She managed to visit virtually all subjects of the source population of 85-year-olds. During these visits, she made personal contact and provided oral and written information on the study. After two weeks and after three months, she visited these subjects again to ask them to participate in the study. The Medical Ethical Committee of the Leiden University Medical Center approved the Leiden 85-plus Study, including the approach and informed consent procedures.

Data collection
For all subjects, socio-demographic characteristics such as gender, marital state, and type of housing were available from the municipal registry. Mean income of neighbourhood of residence was used as indication of socio-economic status (SES). Mean income after taxes in the neighbourhood of residence was obtained by postal codes[2]. We classified low-income neighbourhoods as those with an income below the median.
During the main interview with participants, disability in activities of daily living (ADL) was measured with the Groningen Activity Restriction Scale (GARS)[25]. For participants with severe cognitive impairment, information was obtained from a responsible person. Cognitive function was assessed with the Mini-Mental State Examination (MMSE). Severe cognitive impairment was defined as an MMSE score of 18 points or lower[27]. In participants without severe cognitive impairment, depressive symptoms were measured with the 15-item Geriatric Depression Scale (GDS-15)[27]. At the end of the visit to non-participants as well as to all participants, the research nurse recorded her impression of the subject's daily functioning, cognition, and mood in a standardised questionnaire, using a four-point scale (very good, good, poor, very poor). Validation of the nurse impression about daily functioning, cognition, and mood is presented in the appendix.

All subjects were followed up for all-cause mortality until 1 May 2001. Mortality data were obtained from the municipal registry.

Data analysis
Prevalence estimates of health characteristics by participation status (direct sample, additional input, or non-participants) are presented with 95 percent confidence intervals (95% CI) and are compared using Chi-square tests. Differences in prevalence for socio-demographic and health characteristics between the source population (n = 691) and either the direct sample (n = 511) or the total sample (n = 599)

were assessed by comparing the "true" prevalence of the source population with the calculated confidence interval for the prevalence estimate of both samples. Overall survival was calculated from the 85th anniversary to the date of death or to the date of censoring (1 May 2001). Survival was estimated using the Kaplan-Meier product limit method. Survival by participation status was compared with the log-rank test.

Results

Between 1 September 1997 and 1 September 1999, 705 inhabitants of Leiden reached the age of 85. Fourteen inhabitants died before they could be enrolled in the study and thus 691 subjects were eligible to participate in the study. A total of 511 subjects, the direct sample, participated directly after invitation by phone, resulting in a response rate of 74 percent. After the additional recruitment stage another 88 subjects were included after being personally approached by our research nurse. As a result the total number of participants increased to 599 and the response rate to 87 percent. The remaining 92 subjects refused to participate, of whom 11 subjects refused any contact. Reasons for non-participation were "no interest, no time" (25 percent), "too nervous or anxious" (19 percent), "too tired or ill" (9 percent) or "being against surveys in general" (5 percent). Data from self-report and the nurse's impression were available for 680 subjects (599 participants and 81 non-participants), corresponding to 98 percent of the source population.

Table 1 shows the characteristics of the direct sample (n = 511), the additional input (n = 88) and the non-participants (n = 92). In comparison to the direct sample, subjects from the additional input had poorer health and were more often institutionalised. In contrast, non-participants reported less disability and equal or better health. Subjects from both the additional input and the non-participants reported more often a poor mood.

Differences in characteristics disappeared when we compared the direct sample (n = 511, response 74 percent) and the total sample after the additional recruitment stage (n = 599, response 87 percent) with the source population (n = 691), table 2. Socio-demographic and health characteristics in the source population did not differ from the estimates of these characteristics within the direct and the total sample, except the estimated prevalence of a poor mood.

Figure 1 shows survival by participation status. The 88 subjects who entered the study after the additional recruitment stage had a significantly lower survival compared to the 511 subjects who were directly included (p = 0.04). Survival of the 92 non-participants did not differ from the survival of the direct sample. After inclusion of the additional input with the direct sample (resulting in the total sample) survival functions overlapped (figure 2). Survival of the direct sample as well as survival of the total sample was equal with survival of the source population.

Table 1 *Prevalence estimates and 95% confidence intervals of characteristics by participation status, Leiden 85-plus Study (1997-1999)*

	Direct sample (n=511)			Additional input (n=88)			Non-participants (n=92)		
	n	%	95% CI	n	%	95% CI	n	%	95% CI
Socio-demographics									
Women	335	66%	60-71	62	70%	61-80	72	78 %	70-87*
Institutionalised	82	16%	13-19	26	30%	20-39†	10	11 %	5-17
Widowed	292	57%	53-61	53	60%	50-70	64	70 %	60-79
Low SES	185	36%	32-40	32	36%	26-46	37	40 %	30-50
Self report‡									
Difficulties ADL	208	41%	37-45	45	53%	41-62	20	25 %	16-34†
Poor Health	136	27%	23-31	34	40%	30-51*	17	21%	12-30
Not satisfied	86	17%	14-21	18	24%	14-34	11	14%	6-21
Nurse's Impression‡									
Poor daily functioning	233	46%	41-50	53	60%	50-70*	36	44%	34-55
Poor cognition	183	36%	32-40	53	60%	50-70†	29	36%	25-46
Poor mood	20	4%	2-6	12	14%	7-21†	23	28%	19-38†

** P<0.05 for difference compared with direct sample using Chi square test. † P<0.005 for difference compared with direct sample using Chi square test. ‡ Prevalence estimates are based on total numbers after correction for missing data (11 non-participants refused any contact); When additional input and non-participants were combined, no significant difference with direct sample existed, except for mood.*

Table 2 *Prevalence estimates and 95% confidence intervals of characteristics of the source population and both samples of participants, Leiden 85-plus Study (1997-1999)*

	Source population (n=691)		Direct sample†, 74% response (n=511)			Total sample, 87 % response (n=599)		
	n	%	n	%	95% CI	n	%	95% CI
Socio-demographics								
Women	463	67%	335	66%	61-70	397	66%	62-70
Institutionalised	118	17%	82	16%	13-19	108	18%	15-21
Widowed	403	58%	292	57%	53-61	345	58%	54-62
Low SES	251	36%	185	36%	32-40	217	36%	32-40
Self report‡								
Difficulties ADL	273	40%	208	41%	37-45	253	42%	38-46
Poor Health	187	28%	136	27%	23-31	170	29%	25-33
Not satisfied	115	18%	86	17%	14-21	104	18%	15-21
Nurse's Impression‡								
Poor daily functioning	322	47%	233	46%	41-50	286	48%	44-52
Poor cognition	265	39%	183	36%	32-40	236	39%	35-43
Poor mood	55	8%*	20	4%	2-6	32	5%	4-7

**Significant difference, "true" estimate beyond 95% confidence interval of the estimates from direct and total sample. † Note that subjects of the direct sample are also part of the total sample. ‡ Prevalence estimates are based on total numbers after correction for missing data (11 non-participants refused any contact).*

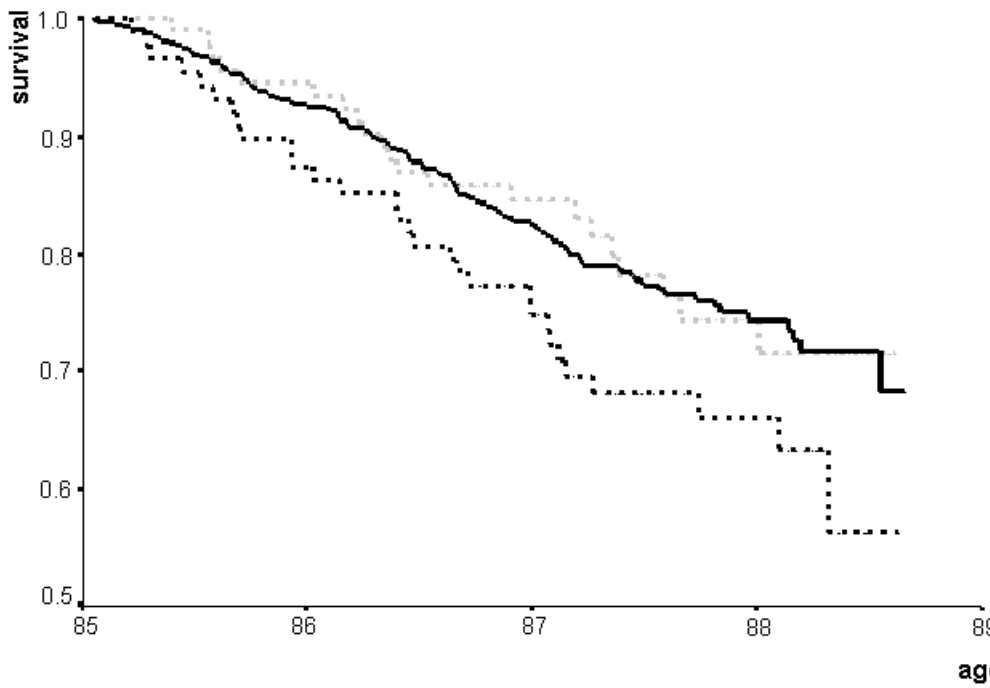

Figure 1 *Cumulative survival for subgroups from the source population.*
Kaplan-Meier estimates of cumulative survival in the participants of the direct sample (n = 511) (continuous line), the additional input (n = 88) (black dotted line) and the non-participants (n = 92) (grey dotted line), Leiden 85-plus Study (1997-1999).

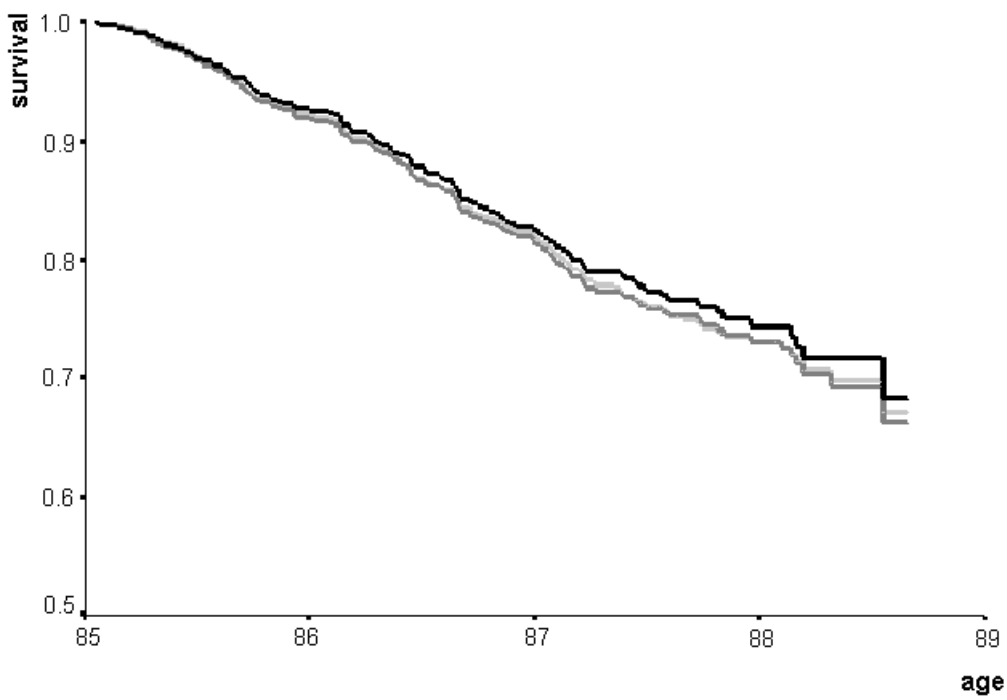

Figure 2 *Cumulative survival for the source population and both samples of participants Kaplan-Meier estimates of cumulative survival in the direct sample (n = 511) (black line), the total sample (n = 599) (dark-grey line) and in the source population (n = 691) (light-grey line), the Leiden 85-plus Study (1997-1999).*

Discussion

The design of our study in which virtually all subjects from the source population were visited at their place of residence, gave us the unique opportunity to compare characteristics of subjects from different samples of participants with all the subjects from the source population, including the non-participants. We tested the hypothesis that the additional effort to increase the response rate would diminish selection bias. We found that the direct sample with a response rate of 74 percent was representative for the source population on baseline characteristics and mortality. With the additional recruitment stage we included frail subjects as shown by a lower survival rate. However, the total sample with a response rate of 87 percent remained representative of the source population. We found that the additional effort to increase the response rate from 74 to 87 percent did not necessarily prevent selection bias. On the contrary, we found that selection bias might have been induced by this effort.

Using data from the Leiden 85-plus Study, we showed that after achieved a representative direct sample with a moderate high response rate, the additional input was a selection of more frail elderly. We used rather crude outcome measures to compare the samples not only on demographic, but also on disabilities, health, and well-being. Using more sensitive measures would not have altered our conclusions that in this population of oldest old additional effort was not necessary to prevent selection bias.

Few studies have mentioned the representativeness of a first wave of recruitment[12] and the possibility of selective additional input[10,28]. Most studies find that particularly frail elderly participate less often in health surveys. Non-participants are described as having a higher age, lower social economic status, lower health status, more depressive mood, lower cognition, and higher morbidity and mortality[1,7,19,29,30]. We found that non-participants were more often depressed but on other characteristics had equal or better health. One could argue that the nurse impression of the mood of non-participants was biased through disappointment and that validation of this impression was done in participants only. However, the high prevalence of poor mood in non-participants is supported by a high proportion (19 percent) of the non-participants who reported depressive symptoms like being too nervous or anxious as the reason for not participating. Moreover, the finding of equal or better health of non-participants might be biased by socially desirable answers[36], since non-participants may have used good health as a reason to support their decision not to participate.

We invested much time and effort in obtaining a very high response rate. The high response rate in our study was due to the personal approach, but other factors of our study design also contributed. Due to the wide publicity our study received, inhabitants of the municipality of Leiden anticipated their 85th birthday letter and felt privileged to belong to the "oldest old". Other factors that might have increased the direct response rate were the involvement of medical staff and nurses instead of lay interviewers, face-to-face interviews at the place of residence[9,10], and an oral informed consent. We think it is conceivable that subjects perceive a written informed consent as a binding contract and therefore refuse to sign anything[31]. The self-confidence and skill in using appropriate heuristics of our expert research

nurse[10], her knowledge of the local situation, and her experience in home care for the elderly were very helpful in achieving a high additional response. Also the enthusiasm of a small research team, where a refusal was felt as a failure might have contributed to the high response rate.

Our design was very time consuming and enabled us to obtain essential information from virtually all eligible subjects. Moreover, survival also appeared to be a good measure to compare subjects by participation status. Mortality is an unbiased outcome that is easily available for both participants and non-participants. Survival rates gave us insight into health differences and comparability of the different samples[16,32,33]. We therefore suggest a wider use of survival as a measure to compare the study population with the source population.

In conclusion, the approach of eligible subjects and the inclusion procedure of population studies are the crux of valid research. We demonstrated that an additional intensive and personal approach is rewarding for drawing more frail elderly subjects into a study. However, this effort will only diminish selection bias if the frail elderly are underrepresented in the direct sample. With an appropriate and conscientious approach the direct sample might already be representative, making additional efforts to increase the response rate to very high levels not necessary.

Appendix

The research nurse recorded her impression on a subject's daily functioning, cognition and mood in a standardised questionnaire at the end of the visit. We compared the scores from these four-point scales (very good, good, poor, very poor) with the scores of the corresponding validated questionnaires on daily functioning, cognition and mood from the main interview as assessed by another member of the medical staff. As the distributions of data were skewed to the left, groups were compared with non-parametric tests that do not assume an underlying normal distribution of the data. As the non-parametric equivalent of the one-way ANOVA procedure, we used the Jonckheere-Terpstra test to determine the p-value for trend between the scores of the questionnaires and the four categories (very good, good, poor, very poor) of the nurse's impression.Results are shown in figures 3A,3B and 3C. The median score for each validated measure showed a gradual and significant ($p < 0.001$) decline or rise over the four categories of the corresponding nurse's impression.

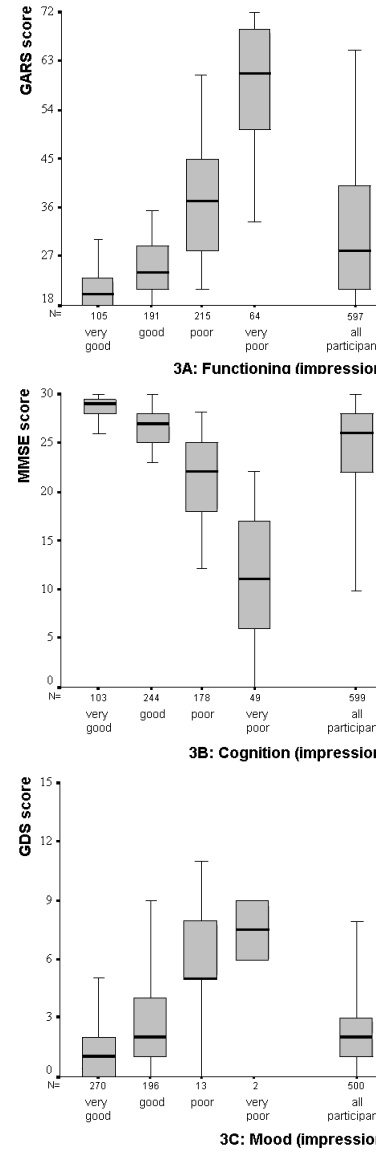

Figure 3 *Comparison between nurse's impression and validated questionnaires*
Comparison of the nurse's impression about daily functioning (A), cognition (B) and mood (C) with test scores on corresponding validated questionnaires within the total sample of participants in the Leiden 85-plus Study (1997-1999). The boxplots show the median (thick line), interquartile range (box) and all values within 5th and 95th percentile.

References

1 Benfante R, Reed D, MacLean C, Kagan A. Response bias in the Honolulu Heart Program. Am J Epidemiol 1989; 130:1088-1100.
2 Reijneveld SA, Stronks K. The impact of response bias on estimates of health care utilization in a metropolitan area: the use of administrative data. Int J Epidemiol 1999; 28:1134-1140.
3 van den Brandt PA, Goldbohm RA, van ‘, V, Volovics A, Hermus RJ, Sturmans F. A large-scale prospective cohort study on diet and cancer in The Netherlands. J Clin Epidemiol 1990; 43:285-295.
4 Deeg DJ, van Tilburg T, Smit JH, de Leeuw ED. Attrition in the Longitudinal Aging Study Amsterdam. The effect of differential inclusion in side studies. J Clin Epidemiol 2002; 55:319-328.
6 Rothman KJ, Greenland S. Precision and validity in epidemiologic studies. In: Rothman KJ, Greenland S, editors. Modern epidemiology. Philadelphia: Lippincott-Raven Publishers, 1998: 115-134.
7 Thompson MG, Heller K, Rody CA. Recruitment challenges in studying late-life depression: do community samples adequately represent depressed older adults? Psychol Aging 1994; 9:121-125.
8 Riedel-Heller SG, Schork A, Matschinger H, Angermeyer MC. Recruitment procedures and their impact on the prevalence of dementia. Results from the Leipzig Longitudinal Study of the Aged (LEILA75+). Neuroepidemiology 2000; 19:130-140.
9 Simonsick EM, Maffeo CE, Rogers SK et al. Methodology and feasibility of a home-based examination in disabled older women: the Women’s Health and Aging Study. J Gerontol A Biol Sci Med Sci 1997; 52:M264-M274.
10 Kessler RC, Little RJ, Groves RM. Advances in strategies for minimizing and adjusting for survey nonresponse. Epidemiol Rev 1995; 17:192-204.
11 Norton MC, Breitner JC, Welsh KA, Wyse BW. Characteristics of nonresponders in a community survey of the elderly. J Am Geriatr Soc 1994; 42:1252-1256.
12 Siemiatycki J, Campbell S. Nonresponse bias and early versus all responders in mail and telephone surveys. Am J Epidemiol 1984; 120:291-301.
14 Cornoni-Huntley J, Ostfeld AM, Taylor JO et al. Established populations for epidemiologic studies of the elderly: study design and methodology. Aging (Milano) 1993; 5:27-37.
15 Dartigues JF, Gagnon M, Barberger-Gateau P et al. The Paquid epidemiological program on brain ageing. Neuroepidemiology 1992; 11 Suppl 1:14-18.
16 Donald IP, Bulpitt CJ. The Gloucestershire Longitudinal Study of Disability: outcomes in nonresponders, responders, and subsequent defaulters. J Clin Epidemiol 1998; 51:1305-1310.
17 Fratiglioni L, Viitanen M, Backman L, Sandman PO, Winblad B. Occurrence of dementia in advanced age: the study design of the Kungsholmen Project. Neuroepidemiology 1992; 11 Suppl 1:29-36.
18 Gordon T, Moore FE, Shurtleff D, et al. Some methodological problems in the long-term study of cardiovascular disease: observations on The Framingham Study. J Chronic Dis 1959; 10:186-206.
19 Hoeymans N, Feskens EJ, van den Bos GA, Kromhout D. Non-response bias in a study of cardiovascular diseases, functional status and self-rated health among elderly men. Age Ageing 1998; 27:35-40.
20 Ott A, Breteler MM, van Harskamp F et al. Prevalence of Alzheimer’s disease and vascular dementia: association with education. The Rotterdam study. BMJ 1995; 310:970-973.
21 Riedel-Heller SG, Busse A, Angermeyer MC. Are cognitively impaired individuals adequately represented in community surveys? Recruitment challenges and strategies to facilitate participation in community surveys of older adults. A review. Eur J Epidemiol 2000; 16:827-835.
22 Rodgers WL, Herzog AR. Collecting data about the oldest old: problems and procedures. In: Suzman RM, Willis DP, Manton KG, editors. The Oldest Old. Oxford: Oxford University Press, 1992: 135-156.
23 Smit JH, de Vries MZ. Procedures and Results of the Fieldwork. In: Deeg DJH, Westendorp-de Seriere M, editors. Autonomy and well-being in the aging population. Report from the Longitudinal Aging Study Amsterdam 1992-1993. Amsterdam: VU University Press, 1994: 7-13.
24 van Tilburg TG, Dykstra PA, Liefbroer AC, Broese van Groenou MI. Sourcebook of Living Arrangements and Social Networks of Older Adults in the Netherlands. 1995. Amsterdam, Faculty of Social Cultural Sciences.
25 Kempen GI, Miedema I, Ormel J, Molenaar W. The assessment of disability with the Groningen Activity Restriction Scale. Conceptual framework and psychometric properties. Soc Sci Med 1996; 43:1601-1610.
26 Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. J Am

Geriatr Soc 1992; 40:922-935.
27 Yesavage JA, Brink TL, Rose TL et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res 1982; 17:37-49.
28 Brenner H. Alternative approaches for estimating prevalence in epidemiologic surveys with two waves of respondents. Am J Epidemiol 1995; 142:1236-1245.
29 Bisgard KM, Folsom AR, Hong CP, Sellers TA. Mortality and cancer rates in nonrespondents to a prospective study of older women: 5-year follow-up. Am J Epidemiol 1994; 139:990-1000.
30 Launer LJ, Wind AW, Deeg DJ. Nonresponse pattern and bias in a community-based cross-sectional study of cognitive functioning among the elderly. Am J Epidemiol 1994; 139:803-812.
31 Carstensen LL, Cone JD. Social desirability and the measurement of psychological well-being in elderly persons. J Gerontol 1983; 38:713-715.
32 Hara M, Sasaki S, Sobue T, Yamamoto S, Tsugane S. Comparison of cause-specific mortality between respondents and nonrespondents in a population-based prospective study: ten-year follow- up of JPHC Study Cohort I. Japan Public Health Center. J Clin Epidemiol 2002; 55:150-156.
33 Heilbrun LK, Nomura A, Stemmermann GN. The effects of non-response in a prospective study of cancer: 15-year follow-up. Int J Epidemiol 1991; 20:328-338.