Caring for vulnerable older people who live in the community

About effective interventions and effective methods for their evaluation

ISBN/EAN: 978-90-9022666-8

Layout and cover: van Druenen Ontwerp & dtp, Veldhoven, the Netherlands

Production: PrintPartners Ipskamp, Enschede, the Netherlands

The work described in this thesis was funded by the Netherlands Organization for Health Research and Development (ZonMw), grant 014-91-055, and Radboud University Nijmegen Medical Centre, and was conducted at the Department of Geriatric Medicine / Nijmegen Alzheimer Centre

Caring for vulnerable older people who live in the community

About effective interventions and effective methods for their evaluation

Een wetenschappelijke proeve op het gebied van
de Medische Wetenschappen

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann
volgens besluit van het College van Decanen
in het openbaar te verdedigen op maandag 28 januari 2008
om 13.30 uur precies door
René Johannes Franciscus Melis
geboren op 5 augustus 1972
te Venhorst

Promotores:	prof. dr. M.G.M. Olde Rikkert
	prof. dr. T. van Achterberg
Copromotores:	dr. G.F. Borm
	dr. E.H. van de Lisdonk
Manuscriptcommissie:	prof. dr. A.L.M. Verbeek (voorzitter)
	prof. dr. G.I.J.M. Kempen (Universiteit Maastricht)
	prof. dr. C. van Weel

Contents

Chapter 1

Introduction

Context of this thesis

A strong call can be heard for a major reform of health care for older adults as we know it today [1]. Several steps along this path have already been taken [2-7]. Among the many reasons that signal the need for this reform, some are of special interest in the context of this thesis: population ageing, an increasing focus on continuing participation, the presence of co-morbidity and vulnerability in a substantial part of the older population, and a scarcity of integration and quality of care. Some of the possible solutions and related problems will be briefly summarised. These considerations ground the choices we have made, when we developed our study model: a model of nurse home visits for vulnerable community-dwelling older people to evaluate aspects of effectiveness, cost-effectiveness, and caregiver burden. A description of the features of our study model will end this analysis of the context of this thesis. We need to have effective methods to study the effects of different types of health care for vulnerable older people. The second part of this context analysis describes which problems were important in the context of our study.

Defining the target population

Before describing the background against which this thesis is set, it is important to identify the target population. The subjects who this thesis focuses on are primarily vulnerable or frail older people. Despite discussions one may have on the exact definition of vulnerability, vulnerable older persons are a distinct group of older adults, with many features that distinguish them from other older adults, as will be explained later. These people can be vulnerable due to (a combination of) multiple chronic conditions, disabilities, or frailty. The compensation capacities of these persons are diminished, and this implies that relatively small changes in or around the patient can have large consequences for the individual's functioning [8]. This thesis further restricts the target population to those frail older persons who still live independently: either in their own homes or in a retirement home.

Reasons for rethinking health care for older adults

The arguments driving health care reform can be grouped into three clusters. These are societal factors, characteristics of older persons, and factors that are inherently related to the way health care is currently conducted.

Societal changes: population ageing and an increasing focus on continuing participation

The first societal development to be mentioned is population ageing. In the next 50 years western societies will face an important demographic change. By the year 2050 16.1% of the world population will be aged 65 and over, and 4.3% will be aged 80 and over. This means more than a doubling for people 65 years and older, and even a threefold increase for octogenarians, compared to 2005 [9]. In particular, the rising proportion of octogenarians will "contribute to the large subset of older adults that are highly vulnerable to adverse health outcomes, including disability, dependency, and death" [10]. Despite some publicists minimising the consequences [11, 12], it is a generally held view that this demographic transition will have a major impact on health care needs and will thus also affect health care delivery. The basic idea behind the necessary change of our health care systems is that we have neither the workforce nor the resources to carry on as usual [13].

Another important societal development is the increasing attention which is given to continuing participation. In more or less explicit terms, this striving for continuing participation can be recognised in many health care reforms. For instance, in the new Dutch Social Support Act (Wet Maatschappelijke Ondersteuning) that took effect from 1 January 2007, this right is translated into a compensation principle: our (local) authorities are obliged to compensate for limitations their citizens may experience in functional abilities and participation in society [5]. The World Health Organization has adopted the term "active ageing" to express its ambitions with respect our ageing societies [3]. In this expression, the word "active" refers to continuing participation in social, economic, cultural, spiritual, and civic affairs. This striving for continuing participation has – more indirectly – also influenced our approaches to health care delivery, think for instance of the many primary care and intermediate care alternatives to hospital care and long term care (care follows the patient). Partly, this development is the consequence of our policy to let people themselves decide where they prefer to live and to be cared for, which is very often their own home [14]. However, based on the principle of continuing participation, policy makers also make an appeal to a person's accountability out of efficiency considerations. As such, continuing participation is not only a granted right; it is also an obligation society lays on its citizens.

Older people: growing numbers of co-morbid conditions and vulnerability

Two characteristics of the older population which have changed over the years are relevant for this thesis: the presence of co-morbidity and vulnerability. Often these two characteristics coexist in one person. Nevertheless, this thesis will judge their implications for health care provision to older people separately. Co-morbidities are indeed an important indicator of vulnerability; however, also in persons with co-morbidity who cannot be classified as vulnerable, health care provision should take the occurrence of co-morbidity into account. In the Netherlands, around two thirds of persons aged 75 and over have one or more chronic conditions, and more than 40% have two or more chronic conditions [15, 16]. This was not the reality decades ago [17], when health care developed as a system with a strong focus on the treatment of acute illnesses; diseases such as infections, or acute myocardial infarction, from which persons died or were cured.

Another important issue is the large group of older persons who are vulnerable because of a combination of co-morbidities, disabilities, or frailty. Although there are distinctions between the concepts of co-morbidity, disability, and frailty as indicators of vulnerability [18-20], this refinement is not needed in the context of this thesis. Therefore, the terms vulnerability and frailty will be used interchangeably throughout this thesis. Vulnerable older people are not simply adults with a higher age, just as children are not simply "petit" adults (a fact long recognised). Vulnerable older people are a distinct group of older people, who deserve special attention and care, just as our children. The compensation capacities of these persons are diminished, and this implicates that relatively small changes in or around the patient can have large consequences for the individual's functioning [8].

Health care systems: fragmented care leads to substandard care

Two problems in health care stand in the way of adequate health care delivery to the group of vulnerable older people. The first problem is that health care today is very

fragmented as well as focused on treatment of single diseases [21]. We have single-disease management programmes, disease-based clinical practice guidelines, organ-specific specialists, etc. As a result we fail to see the interactions at biological, psychological, and social levels, lack understanding of how to deal with diseases in the light of co-morbidity [22], and are blind for the vulnerability of the older person in front of us. Individuals cannot rely on an integrated system of facilities tailored to their needs; instead they have to find their way through fragmented settings of care. This approach was and is very successful in fairly fit, independent, assertive, (cognitively) healthy adults experiencing one single problem, but can be a disaster for the vulnerable, cognitively impaired older adult with co-morbidity. A recent study has shown how binding to the several single-disease guidelines may lead to unintended and undesirable effects in older persons with several co-morbidities [23].

The quality of care for people aged 65 and over has been studied previously [23-25]. The results leave little room for discussion: "care for vulnerable elders falls short of acceptable levels for a wide variety of conditions. Care for geriatric conditions is much less optimal than care for general medical conditions" [24]. Several causes can be identified: lack of (dissemination of) knowledge, pessimistic expectations about outcomes, and organisational barriers [10].

Strategies to improve quality of care for vulnerable older people

Several solutions have been proposed at different levels. In many of these strategies one can identify aspects of attitude, knowledge, and structure as a means to provoke changes at a behavioural level in professionals as well as patients.

Change of "attitude" strategies are for instance strategies that try to increase patient-centredness by increasing the health care professionals' awareness of the importance of patients' involvement as partners in care, or to make persons aware of the fact that they can try themselves – with some help from their family – first, before requesting help from the authorities or professionals. The problem with "attitude" changes is how to enforce them and have them widely implemented in health care. With respect to enforcement, to increase knowledge is an easier target, for instance through a formalised system of continuing medical education, continuing professional development and knowledge translation [26]. The last is an example of a structural change of health care systems.

Patient-centred care: a popular but difficult concept

Of all strategies, the strategies to improve patient-centred care rely the most on changing "attitude", and therefore the problem of enforcement is extremely relevant for this type of health care reform. To what extent professionals perform patient-centred care is highly dependent on their individual judgement of "good care". Patient-centred care still is a poorly understood, rather intuitively defined concept [27] that is not easily accessible for empirical study. As a result, the extend to which of patient-centred care can contribute to a better primary care for vulnerable older people is largely unknown.

Care models for vulnerable older people

Programmes of integrated care focus on the structural level. Integrated care programmes for chronic diseases seem to have positive effects on quality of care [28]. Although Ouwens et al did not look specifically at vulnerable older people, vulnerable older people were

an important target group of the interventions under study in this systematic review.

If we look in more detail at the group of vulnerable older people, we can observe that this group has shown to benefit from inpatient geriatric care models, such as inpatient geriatric evaluation and management at a specialised geriatric ward [29]. However, the effectiveness of in-home or outpatient alternatives is much more controversial. These are models such as preventive home visits, in-home comprehensive geriatric assessment, and intermediate care, which may provide effective care. However, literature is not unequivocal about the question to which type of older persons geriatric primary care should be directed. While some authors exclude the frailest participants, because these persons have too few possibilities for reversibility, others stress the importance of including the frailest [30-35]. Although the above discussion focuses primarily on preventive home visits, because the boundaries between community intervention models such as preventive home visits, in-home geriatric evaluation and management as well as intermediate care are so vague, this discussion is of relevance for all these models. The models that can be gathered under the term community intervention models show much heterogeneity as well considerable overlap [36]. The lack of detailed insight into the content of these care models further complicates comparison [31, 37].

However, we can try to distil elements that may be associated with beneficial patient outcomes [32, 38]. A detailed description these elements is given in the background section of chapter 3, which describes the design of our randomised trial. One element to be mentioned here is the involvement of the general practitioner in these models. Direct involvement of the general practitioner has been identified as a criterion for success [32], also because providers' cooperation is a determinant of patient adherence to programme recommendations [39]. Another element that may be related to increased effectiveness is the use of targeting [32, 38]. Several approaches to targeting are available, such as population screening [40], targeting on the basis of patients' (medical) history [41], or a problem-based approach where incident problems are used to start an intervention.

Costs: does primary care geriatrics provide value for money?

A related issue that deserves our attention is the economics of all the strategies of improvement. Referring to the rising health care costs – a consequence of technological innovations as well as population ageing – the Dutch Council for Public Health and Health Care (Raad voor volksgezondheid en zorg; RVZ) has highlighted the urgency of making choices[42]. This council suggested a budgeting ceiling of 80,000 euros per Quality Adjusted Life Year (QALY), and NICE (National Institute for Health and Clinical Excellence) – an important UK institute that provides information to guide health care technology choices – has stated that their guidance has to "explain, explicitly, reasons for recommending – as cost effective – those interventions with an incremental cost-effectiveness ratio in excess of £20,000 to £30,000 per QALY" [42, 43]. These thresholds are given because they can help making choices between several health care types more objectively. These choices are now made on an ad hoc basis [42]. In this circus of boosting opinions and bold language, the marketing of older people's interests and care has proven to be difficult. Therefore, these threshold values – as they make the decision making more transparent and objective, may be a blessing for geriatric care, but only if information on cost-effectiveness is available. Regrettably, cost-effectiveness data are sparse in the field of geriatric primary care. The background section of chapter 7 – in

which the cost-effectiveness data of our intervention programme are presented – will go into this matter in more detail.

Considerations on the design of DGIP:
why did we design our study model as we did?

The Dutch Geriatric Intervention Programme (DGIP) was a multidisciplinary community intervention model, consisting of house calls for frail older patients. An extensive description of the model can be found in the background section of chapter 2 and the methods sections of chapter 6. The core was a multidisciplinary team intervention, carried out by general practitioner, geriatrician, and specialist geriatric nurse, who had a pivotal role in the intervention.

The reason to have nurses conducting our intervention is they are able to deliver high quality primary care [44], which provides a good complement to the care delivered by physicians. Also, in the field of chronic disease management and coordination of care, nurses are highly effective [44]. Multidisciplinarity itself is related to positive outcomes of geriatric (primary) care [38].

We hypothesised that the effectiveness of community intervention models for frail older people could be enhanced using a problem-based participant selection performed by their general practitioner. General practitioners were asked to initiate the intervention when a geriatric condition arose that required further intervention. This procedure promised continuity of care through better timing and targeting of the intervention, and more engagement of the general practitioner. Moreover, patient selection is no longer dependent on an expensive population screening procedure, which enhances implementation chances. Problems had to relate to cognition, behaviour, mood, mobility, or nutrition, which are important geriatric syndromes.

Research issues in studying geriatric primary care

Along with the compelling need for effective health care for older people, comes the increasing need for effective methods for their study. Several challenges are attached to performing research on complex interventions for vulnerable older people in a primary care setting. How to motivate general practitioners, who have many other tasks and priorities? How to engage vulnerable, cognitively impaired, and sometimes mentally ill persons, in follow-up visits packed with strange, difficult, and often very private questions? How to compare results of research projects, identify gaps in our current knowledge, or critically appraise the ascribed benefits, if there is much disagreement on definitions of the object under study? The three main research challenges of this study are related to the fact that we aimed to study the effects of a complex intervention for frail older people in a primary care setting. Each aspect has individual difficulties attached.

Older people are excluded from research because of a multitude of reasons, reflecting the diversity of the older population: legal incompetence, the presence of co-morbidity, the use of competing medication, etc. [45]. Some important work on improving the methodology of doing research with older people has been done by The Interventions on Frailty Working Group [46]. Important challenges they identified related to eligibility criteria (too strict limits generalisability), consent procedures (related ethical difficulties in the legally incompetent), self reports (may be problematic), attrition, competing morbidity,

mechanism of the intervention (maybe unclear), and outcome measures (improvements in functional status may not translate into well-being and quality of life).

Some of the problems involved in doing research in a primary care setting are rather prosaic: high workloads, financing problems, and sometimes fewer academic interests, or less organised environments than in hospital care [47]. As a result, research may not rank high on the list of priorities. More fundamental difficulties are that primary care works with less selected populations (this may be judged an advantage as well, though still a challenge) with often multimorbidity, and compliance and motivation to participate may be problematic, etc. [47, 48].

The evaluation of complex or pragmatic interventions has difficulties because there is often no complete consensus on the definition, or the exact content of the intervention under study. Both problems harms scientific appraissal of the available evidence. A paper on the methodology of the evaluation of complex interventions by Campbell et al puts it like this: "the evaluation of complex interventions is difficult because of problems of developing, identifying, documenting, and reproducing the intervention" [49].

Currently, the double blind, placebo-controlled, randomised controlled trial is widely accepted as the most reliable tool to study the efficacy of new treatment. However, this method cannot readily be used in the study of complex interventions. For instance, in evaluations of complex interventions patients and the professionals who deliver the interventions hardly ever can be blinded to the treatment they are assigned to receive or provide. Even though patients are randomised, selective drop out mechanisms may cause distortion of the results. For the same reason, contamination may occur. Contamination occurs when some elements of one treatment mix into the other treatment, and vice versa. Some of the problems and difficulties described above were more relevant to our trial than others. Next, the most relevant problems will be discussed.

Loss-to-follow-up

Loss-to-follow-up can be an important threat to the validity and generalisability of our study. Firstly, because we aimed to study frail older people who probably have an increased susceptibility to dropping out because of their vulnerability. Secondly, because we studied a complex intervention, the delivery of which cannot be blinded, selective drop outs may occur. We took several precautionary measures to deal with the problems of potential dropping out. While we accounted for some loss-to-follow-up in our sample size considerations, our primary goal was to prevent the occurrence of loss-to-follow-up as much as possible. We did so, because this is the best safeguard against differential attrition after all.

In advance we decided on a set of outcome measures which were not too burdensome for vulnerable persons, and identified an alternative strategy to acquire data for patients who felt this was too troublesome after all. In order to minimise the problems of unblinding of the assessor, we made sure that the primary outcome measures were collected using a written questionnaire that was filled out before each follow-up visit (if necessary with help from a relative).

Length of follow-up

The length of follow-up is another issue we considered. The following elements were taken into account.

Generally, longer follow-up periods are of more relevance than shorter time periods (if one can show that an intervention has longlasting effects, this intervention is probably more clinically relevant than an intervention with only short term effects). However, the frail population we included had many competing risks, and a fairly limited life span. Improving quality of life in the short term therefore is important as well. These effects would have been missed if we had only looked at the longer term (by that time the effects of the intervention might already have been diminished due to other competing risks). One can of course decide to establish the effect at multiple time points, during a longer follow-up period. However, this might again become too burdensome for our frail subjects. Also, one still has to decide which time point is of most clinical relevance: earlier, or later during follow-up. Given the above considerations, we judged the short term effects to be of prime importance. Therefore, we decided to use fairly short term follow-up measurements of three and six months.

Outcome measures

Interventions directed at the very old and vulnerable should focus primarily on increasing autonomy and quality of life. For successful ageing, adding life to years seems of more relevance than adding years to life. Therefore we gave priority to health-related quality of life measures over survival measures. Unfortunately, it was difficult to find valid outcome measures that were applicable to a very diverse, yet very frail population of older adults, and were also well-known. Another hurdle was to find information on the clinical relevance of effects and changes, especially for a group of vulnerable older adults [50]. Another area that requires further development is the area of responder analysis: when to call a treatment a success [51]. This information provides an important addition – yet no alternative – to numeric outcome measures, for doctors who want to understand – and value the clinical relevance of – a difference.

Unable to double blind: threats of selection bias, recruitment problems, and contamination

A specific issue we were confronted with concerned an important methodological dilemma: how to prevent the occurrence of contamination bias in a randomised controlled trial of an intervention that cannot be administered blindly (as with many complex interventions), while at the same time ensuring comparability of intervention and control group? Contamination is dilution of contrast between study arms, because the treatment alternatives somehow get mixed up. In our study, the general practitioners' exposure to the intervention resulting from their participation could lead to contamination of control patients and thus introduce contamination bias, when their patients had been randomised individually. However, the widely-used solution to avoid contamination bias, cluster randomisation – randomising all patients of one general practitioner to the same study arm –, risked selection bias and recruitment problems. No ready-made solutions to this dilemma were available. We solved the problem by combining cluster and individual randomisation in a new randomisation method called pseudo cluster randomisation.

Research aims

The overall aim of this thesis is to describe the effects of a new multidisciplinary model of nurse-led, house call based health care provision for independently living vulnerable older people on patient, caregiver, and cost-effectiveness outcomes.

The research questions are:

What are the effects of the Dutch Geriatric Intervention Programme (DGIP) compared to usual care in improving health-related quality of life in independently living elderly persons, on caregiver burden, on health care use of frail older people, and on patient survival? What is the cost-effectiveness compared to usual care from a health care system's point of view? With respect to the dilemma of contamination versus selection bias, we described the method of pseudo cluster randomisation, the sustainability of the premises underlying the use of pseudo cluster randomisation, and the effect of pseudo cluster randomisation on study recruitment and validity.

Outline of this thesis

The main findings of our studies are presented in chapter 6, 7, and 8. Chapter 6 presents the results for the patient outcomes health-related quality of care, and survival. Chapter 7 describes the effects on health care utilisation and cost-effectiveness. Chapter 8 provides a detailed overview of the effect on caregiver burden. Chapter 2 applies the example of intermediate care to go in more detail into the matter of disagreeing definitions and how this hampers scientific appraisal of the available knowledge. Chapter 3 is a detailed description of the study design of our main study. Chapters 4 and 5 go in more detail into pseudo cluster randomisation; chapter 4 describes the application of the method, chapter 5 the results of the evaluation. The final chapter 9 summarises the results and puts them in a broader methodological, scientific, and societal context.

References

1. Website European Observatory on Health Systems and Policies. http://www euro who int/observatory 2006 September 11;Available at: URL: http://www.euro.who.int/observatory. Accessed September 22, 2006.
2. The National Service Framework for older people. London: Department of Health; 2001.
3. WHO. Active ageing: a policy framework. Geneva: World Health Organisation; 2002.
4. Primary care in the driver's seat? Organizational reform in European primary care. Maidenhead, Berkshire: Open University Press; 2006.
5. Wet Maatschappelijke Ondersteuning (Social Support Act), Wet Maatschappelijke Ondersteuning (Social Support Act), (2006).
6. Lang zullen we leven (Rapport Themacommisie Ouderenbeleid Tweede Kamer der Staten Generaal). 's Gravenhage: Sdu Uitgevers; 2005.
7. 2005 White House Conference on Aging. 2005 White House Conference on Aging. The Booming Dynamics of Aging: From Awareness to Action. Report to the President and the Congress. Bethesda MD: 2005 White House Conference on Aging; 2006.
8. Slaets JP. [The importance of the 'frailty-concept' in geriatric assessment]. Tijdschr Gerontol Geriatr 1998 December;29(6):276-8.
9. United Nations. World population prospects; the 2004 revision analytical report. New York, NY: Population Division, Department of Economic and Social Affairs, United Nations (UN); 2004.
10. Fried LP. Establishing benchmarks for quality care for an aging population: caring for vulnerable older adults. Ann Intern Med 2003 November 4;139(9):784-6.
11. Gee EM. Misconceptions and misapprehensions about population ageing. Int J Epidemiol 2002 August;31(4):750-3.
12. Denton FT, Gafni A, Spencer BG. Requirements for physicians in 2030: why population aging matters less than you may think. CMAJ 2003 June 10;168(12):1545-7.
13. Raad voor Volksgezondheid en zorg. Arbeidsmarkt en zorgvraag (The labour market and demand for care). Zoetermeer: Raad voor Volksgezondheid en Zorg; 2006.
14. Salvage AV, Jones DA, Vetter NJ. Options of people aged over 75 years on private ad local authority residential care. Age Ageing 1989;18:380-6.
15. van Lindert H, Droomers M, Westert GP. Tweede Nationale Studie naar ziekten en verrichtingen in de

huisartspraktijk. Een kwestie van verschil: verschillen in zelfgerapporteerde leefstijl, gezondheid en zorggebruik. [Second Dutch National Study on diseases and interventions in GP practice. A matter of difference: differences in selfreported life style, health, and care utilisation]. Utrecht, Bilthoven: NIVEL, Rijksinstituut voor Volksgezondheid en Milieu; 2004. Report No.: 3.
16. StatLine. Voorburg: Statistics Netherlands; 2005.
17. Deeg DJ, Kriegsman DMW, van Zonneveld RJ. Prevalentie van vier chronische ziekten en hun samenhang met gezondheidsbeperkingen bij ouderen in Nederland, 1956-1993 (Prevalence of four chronic diseases and their association with health limitations in older persons in The Netherlands, 1956-1993). Tijdschr Soc Gezondheidsz 1994;72(8):434-41.
18. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. J Gerontol A Biol Sci Med Sci 2004 March;59(3):255-63.
19. Fried LP, Tangen CM, Walston J et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 March;56(3):M146-M156.
20. Markle-Reid M, Browne G. Conceptualizations of frailty in relation to older adults. J Adv Nurs 2003 October;44(1):58-68.
21. Tinetti ME, Fried T. The end of the disease era. Am J Med 2004 February 1;116(3):179-85.
22. Tinetti ME, Bogardus ST, Jr., Agostini JV. Potential pitfalls of disease-specific guidelines for patients with multiple conditions. N Engl J Med 2004 December 30;351(27):2870-4.
23. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW. Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance. JAMA 2005 August 10;294(6):716-24.
24. Wenger NS, Solomon DH, Roth CP et al. The quality of medical care provided to vulnerable community-dwelling older patients. Ann Intern Med 2003 November 4;139(9):740-7.
25. Jencks SF, Cuerdon T, Burwen DR et al. Quality of medical care delivered to Medicare beneficiaries: A profile at state and national levels. JAMA 2000 October 4;284(13):1670-6.
26. Davis D, Evans M, Jadad A et al. The case for knowledge translation: shortening the journey from evidence to effect. BMJ 2003 July 5;327(7405):33-5.
27. Stewart M. Towards a global definition of patient centred care. BMJ 2001 February 24;322(7284):444-5.
28. Ouwens M, Wollersheim H, Hermens R, Hulscher M, Grol R. Integrated care programmes for chronically ill patients: a review of systematic reviews. Int J Qual Health Care 2005 April;17(2):141-6.
29. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. Lancet 1993 October 23;342(8878):1032-6.
30. Egger M. Commentary: When, where, and why do preventive home visits work? BMJ 2001;323(7315):724-5.
31. Elkan R, Kendrick D, Dewey M et al. Effectiveness of home based support for older people: systematic review and meta-analysis. BMJ 2001 September 29;323(7315):719-25.
32. Rubenstein LZ, Stuck AE. Preventive home visits for older people: defining criteria for success. Age Ageing 2001 March;30(2):107-9.
33. van Rossum E, Frederiks CM, Philipsen H, Portengen K, Wiskerke J, Knipschild P. Effects of preventive home visits to elderly people. BMJ 1993 July 3;307(6895):27-32.
34. Gill TM, Baker DI, Gottschalk M, Peduzzi PN, Allore H, Byers A. A program to prevent functional decline in physically frail, elderly persons who live at home. N Engl J Med 2002 October 3;347(14):1068-74.
35. Stuck AE, Beck JC, Egger M. Preventing disability in elderly people. Lancet 2004 November 6;364(9446):1641-2.
36. Parker G, Bhakta P, Katbamna S et al. Best place of care for older people after acute and during subacute illness: a systematic review. J Health Serv Res Policy 2000 July;5(3):176-89.
37. Walsh B, Steiner A, Warr J, Sheron L, Pickering R. Nurse-led inpatient care: opening the 'black box'. Int J Nurs Stud 2003 March;40(3):307-19.
38. Wieland D. The effectiveness and costs of comprehensive geriatric evaluation and management. Crit Rev Oncol Hematol 2003 November;48(2):227-37.
39. Bula CJ, Alessi CA, Aronow HU et al. Community physicians' cooperation with a program of in-home comprehensive geriatric assessment. J Am Geriatr Soc 1995 September;43(9):1016-20.
40. Wieland D, Rubenstein LZ. What do we know about patient targeting in geriatric evaluation and management (GEM) programs? Aging (Milano) 1996 October;8(5):297-310.
41. Implementing the Evercare programme. Interim report. London: United Health Group; 2004.
42. Raad voor Volksgezondheid en zorg. Zinnige en duurzame zorg [sensible and sustainable care]. Den Haag: Raad voor Volksgezondheid en Zorg; 2006.
43. National Institute for Health and Clinical Excellence. Social Value judgements, principles for the development of NICE guidance. London: National Institute for Health and Clinical Excellence; 2005.
44. Sibbald B, Laurant MG, Reeves D. Advanced nurse roles in UK primary care. Med J Aust 2006 July 3;185(1):10-2.
45. McMurdo ME, Witham MD, Gillespie ND. Including older people in clinical research. BMJ 2005 November 5;331(7524):1036-7.
46. Ferrucci L, Guralnik JM, Studenski S, Fried LP, Cutler GB, Jr., Walston JD. Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. J Am Geriatr Soc 2004 April;52(4):625-34.
47. Ward E, King M, Lloyd M, Bower P, Friedli K. Conducting randomized trials in general practice: methodological and practical issues. Br J Gen Pract 1999 November;49(448):919-22.
48. Thomas P. The research needs of primary care. BMJ 2000 July 1;321(7252):2-3.

49. Campbell M, Fitzpatrick R, Haines A et al. Framework for design and evaluation of complex interventions to improve health. BMJ 2000 September 16;321(7262):694-6.
50. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999 February;15(2):141-55.
51. Oppenheimer L, Kher U. The impact of measurement error on the comparison of two treatments using a responder analysis. Stat Med 1999 August 30;18(16):2177-88.

Chapter 2

What is intermediate care?
An international consensus on what constitutes intermediate care is needed

R.J.F. Melis, S.G. Parker, M.I.J. van Eijken, M.G.M. Olde Rikkert

BMJ 2004;329:360–1

Intermediate care is an emerging concept in health care, which may offer attractive alternatives to hospital care for elderly patients. As little scientific evidence exists on the benefits of intermediate care, research is especially important.[1-3] A prerequisite for research is agreement on the definition of a concept, which is lacking for intermediate care. The term intermediate care is often used as if its meaning is clear, but it conveys little meaning other than being about care that is "in between". Commonly used definitions of intermediate care do not help much, and several very different definitions are in use. What is needed at the outset is a consensus on what constitutes intermediate care. Until this is agreed on, the concept of intermediate care will remain a mirage and its possibilities unknown.

The term intermediate care was introduced in the United Kingdom's NHS Plan and refined in the national service framework for older people.[4, 5] The concept seems to arise out of a policy imperative, rather than an analysis of the scientific evidence about effective models of care. Objectives such as "promotion of independence" and "prevention of unnecessary hospital admission" were to be achieved through providing a new range of services between hospital and home. Specific targets (for example, the number of service users, prevented admissions) accompanied these objectives. However, no particular models of service delivery were defined.

Professional statements of good practice followed the political decision that reconfiguration of the health service would include investment in intermediate forms of care. The British Geriatrics Society listed three definitions in its statement on intermediate care.[1] The broadest definition is the one shared with the Royal College of Physicians of London, according to which intermediate care is delivered by those health services that do not require the resources of a general hospital but are beyond the scope of the traditional primary care team.[2]

Recently, Andrea Steiner published as many as eight definitions of intermediate care.[6] Five of them (partly) focused on facilitating the transition from hospital to home. Other aims include avoidance of admission and improvement of pre-acute and post-acute care. A systematic review on the best place of care for older people after acute illness concluded that service models were best described in terms of the objectives of care.[7]

Definitions from the databases Medline and CINAHL narrow intermediate care in the direction of nursing home care. For example, in Medline "intermediate care facilities" are institutions that provide health-related care and services to individuals who do not require the degree of care which hospitals or skilled nursing facilities provide, but require care and services above the level of room and board. This probably results from the existence of intermediate care facilities in the United States and Japan, which closely resemble nursing homes. Further difficulty arises because individual authors also use the term intermediate care when describing a less advanced type of intensive care medicine.[8]

This inventory of definitions shows that the term intermediate care currently does not present imply a specific, well defined type of health care (box). This worrying conclusion has important consequences. To compare results of research projects will be difficult if not impossible, as will be identifying gaps in our current knowledge or critically appraising the benefits attributed to intermediate care. These difficulties will only increase because

of the growing popularity of alternatives to hospital inpatient care across Europe and the rest of the world.

box *Definitions of intermediate care*

British Geriatrics Society[1]

- An approach to health care intended to facilitate patients' transitions from illness to recovery, or to prevent their transition from home managed chronic impairment to institution-based dependence, or to help terminally ill people be as comfortable as possible at the end of their lives
- That range of services designed to facilitate transition from hospital to home, and from medical dependence to functional independence, where the objectives of care are not primarily medical, the patients' discharge destination is anticipated, and a clinical outcome of recovery (or restoration of health) is desired
- Those services that do not require the resources of a general hospital, but are beyond the scope of the traditional primary care team. These can include "substitutional care" and "care for people with complex needs"

(The last definition is the same as the one the Royal College of Physicians uses in its statement[2])

Medical subject heading (MeSH)[9]

- Intermediate care facilities are institutions that provide health-related care and services to individuals who do not require the degree of care that hospitals or skilled nursing facilities provide, but because of their physical or mental condition require care and services above the level of room and board

CINAHL subject headings

- Intermediate care (see subacute care) is care provided to acute care patients who are medically stable but too unstable to be treated in alternative healthcare settings such as home, ambulatory, or traditional skilled long term care
- Intermediate care facilities: entered here are materials on nursing home facilities. For care given in a nursing home, see long term care

To deal with this Babel of voices we suggest a formal process to develop a consensus of the key elements of intermediate care. The aim of this debate should not be to arrive at a uniform definition of intermediate care, for our inventory on the definitions of intermediate care has shown that it is impossible to define intermediate care unequivocally at the highest conceptual level. For reasons of simplicity, this debate should be limited to defining intermediate care for the purpose of scientific appraisal. It would also be helpful if bibliographers were able to establish a consensus for terminology, such as medical subheadings.[9] For the time being we believe that intermediate care models can be best classified according to their objectives of care and not by their names. If we do not clearly define key elements of the concept of intermediate care, then it will remain a concept with unfulfilled promise.

References

1. Intermediate care. Guidance for commissioners and providers of health and social care (BGS Compendium Document D4). London: British Geriatrics Society; 2001.
2. Black C, Black D, Alberti G. Intermediate care: statement from the Royal College of Physicians of London. London: The Royal College of Physicians; 2000 Oct 5.
3. Carpenter I, Gladman J, Parker S, Potter J. Clinical and research challenges of intermediate care. *Age Ageing* 2002;31:97-100.
4. The NHS Plan. A plan for investment, a plan for reform. London: Stationary office; 2000.
5. The National Service Framework for older people. London: Department of Health; 2001.
6. Steiner A. Intermediate care--a good thing? *Age Ageing* 2001 August;30 Suppl 3:33-9.
7. Parker G, Bhakta P, Katbamna S et al. Best place of care for older people after acute and during subacute illness: a systematic review. *J Health Serv Res Policy* 2000 July;5(3):176-89.
8. Junker C, Zimmerman JE, Alzola C, Draper EA, Wagner DP. A multicenter description of intermediate-care patients: comparison with ICU low-risk monitor patients. *Chest* 2002 April;121(4):1253-61.
9. US National Library of Medicine. 2003 Medical subject heading, annotated alphabetic list. Springfield, Virginia: National Technical Information Services; 2003.

Chapter 3

The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]

R.J.F. Melis, M.I.J. van Eijken, G.F. Borm, M. Wensing, E. Adang, E.H. van de Lisdonk, T. van Achterberg, M.G.M. Olde Rikkert

BMC Health Services Research 2005;5:65

Background: Because of their complex clinical presentations and needs frail elderly people require another approach than people who age without many complications. Several inpatient geriatric health services have proven effectiveness in frail persons. However, the wish to live independently and policies that promote independent living as an answer to population ageing call for community intervention models for frail elderly people. Maybe models such as preventive home visits, comprehensive geriatric assessment, and intermediate care qualify, but their efficacy is controversial, especially in frail elderly persons living in the community. With the Dutch EASYcare Study Geriatric Intervention Programme (DGIP) we developed a model to study effectiveness of problem-based community intervention models in frail elderly people.
Methods: DGIP is a community intervention model for frail elderly persons where the general practitioner refers elderly patients with a problem in cognition, mood, behaviour, mobility, and nutrition. A geriatric specialist nurse applies a guideline-based intervention with a limited number of follow-up visits. The intervention starts with the application of the EASYcare instrument for geriatric screening. The EASYcare instrument assesses (instrumental) activities of daily life, cognition, mood, and includes a goal setting item. During the intervention the nurse regularly consults the referring general practitioner and a geriatrician. Effects on functional performance (Groningen Activity Restriction Scale), health-related quality of life (MOS-20), and caregiver burden (Zarit Burden Interview) are studied in an observer blinded randomised controlled trial. 151 participants were randomised over two treatment arms – DGIP and regular care – using pseudo cluster randomisation. We are currently performing the follow-up visits. These visits are planned three and six months after inclusion. Process measures and cost measures will be recorded. Intention to treat analyses will focus on post intervention differences between treatment groups.
Discussion: The design of a trial evaluating the effects of a community intervention model for frail elderly people was presented. The problem-based participant selection procedure satisfied; few patients that the general practitioner referred did not meet our eligibility criteria. The use of standard terminology makes detailed insight into the contents of our intervention possible using terminology others can understand well.

Background

In frail elderly persons chronic conditions and loss of function challenge their autonomy. This harms their well-being, and often leads to institutionalisation and high health care costs.

There is much heterogeneity in the degree to which frailty affects older people. While some have many problems, others age successfully [1]. The introduction of the concept of successful ageing voiced a change in our thinking about "age-related" decline [2]. It marked the awareness that we cannot simply regard functional loss and dependency as consequences of the ageing process itself when disease is absent. With this understanding these "age-related" deficits became amenable to intervention. Of course, intervention should take the heterogeneity of the population into account; because of their complex clinical presentations and needs frail persons require another approach than people who age without many complications [3]. Although special services for frail older people have proven effectiveness in the form of inpatient geriatric health services [4], several societal developments ask for community equivalents. People prefer to stay at home, even with considerable disability [5]. Another drive behind the development of community intervention models comes from policies that promote independent living as an answer to the questions raised by population ageing [6]. Possibly, models such as preventive home

visits, in-home comprehensive geriatric assessment, and intermediate care provide effective health services for frail older people in the community. Unfortunately, both the definition and efficacy of these community intervention models remain subject of a vivid debate [7-10]. The debate stems from the fact that the models gathered under these names show much heterogeneity as well as considerable overlap [11]. The lack of detailed insight into the content of these care models further complicates comparison [12, 13]. One of the major issues is the effectiveness of these models in the expanding group of frail older people.

Despite the diversity, from literature we can distil certain elements that are used in many community intervention models. These are elements such as multidimensional and multidisciplinary working, person centred care, participant selection, and treatment adherence. Empirical evidence is available for some of these elements.

In this paper we will briefly summarise this knowledge on multidimensional assessment and management of elderly people in the community. This information grounds the choices we have made in designing a new community intervention model for frail elderly people living at home. Then, we will present the outlines of our intervention model and the design of the randomised trial in which we are currently evaluating the model. At this moment the recruitment period is already completed, and we are performing the follow-up visits. Therefore, in addition to the details of the design, we will highlight some results of the conduct of the recruitment phase of our study.

Evidence on intermediate care models

Most research has been done on preventive home visits and comprehensive geriatric assessment, less scientific knowledge is available for intermediate care models.

The studies that have been evaluating intermediate care focused mainly on the evaluation of intermediate care alternatives (e.g. rapid response teams, hospital at home, early discharge schemes) in direct comparison with hospital care [11, 14]. For most of the models that are not intended as direct alternatives to hospital care (e.g. residential rehabilitation, and community assessment and rehabilitation services) only descriptive data are available [15].

(Evidence-based) elements of community intervention models

Virtually all community intervention models for older people share a similar multidimensional nature covering a variety of medical, psychological, functional, and social domains. As multidimensional working is a ubiquitous feature of these models, it is in itself not thoroughly studied. There are some discussions on which domains are to be included [8].

Both in preventive home visits and comprehensive geriatric assessment it is suggested that models with a multidisciplinary team are more effective than models with a unidisciplinary approach [8, 16]. Effectiveness is also claimed for longer follow-up and more home visits, although a recent trial did not confirm this [16, 17].

Many models provide person centred care. Some even argue that “patient-centred, problem-driven, goal-oriented management” is a “key minimum specification” [16].

Another element that might strengthen the effectiveness of comprehensive geriatric assessment is to secure control over the implementation of the recommendations done in the programme [4]. Models implemented in regular care often do not have complete clinical control over the enforcement of the recommendations following from the programme. In this scenario, it is very important to involve the primary care provider who will be responsible for the implementation of the proposed plan [8]. This is also important because providers' cooperation is a determinant of patient adherence to programme recommendations [18]. It is difficult to change physicians' behaviour and this urges the use of high intensity programmes. Furthermore, programme effectiveness might benefit from stronger emphasis on direct recommendations to participants, and should not rely too much on the uptake of recommendations by the primary care provider [8].

Participant selection is a feature of community intervention models for elderly people that received much attention in literature. This discussion focuses on two matters: participant selection on the basis of age and on the basis of participants' needs. Age as a selection criterion is not discussed much, but causes controversy. Some authors state that home visits are more effective in persons aged 75 and over, compared to younger individuals [19]. One meta-analysis did not find an age effect, and another meta-analysis concluded most benefits are to be expected in the youngest old [13, 20]. Frailty has received much more attention than age with respect to targeting these health services models to those who will benefit most. Most authors agree that too healthy elderly persons should be excluded, because both preventive home visits and comprehensive geriatric assessment are ineffective in these sprightly people [13, 21]. There is more dispute about the effectiveness of these models in frail older persons. While some exclude the frailest participants, because in these persons there are too few possibilities for reversibility, other authors stress the importance of including the frailest [8, 13, 21, 22]. Combining the evidence on the relevance of both age and frailty for participants selection Stuck concludes that health risk appraisal with individual reinforcement is beneficial to healthy persons aged 60 to 75, preventive home visits should focus on independent people aged 75 and over, and that other types of (institutional) services are needed for the frailest [23].

Unfortunately, considered this is true, this conclusion still disregards the population of frail elderly persons living in their own home.

Considerations on designing the Dutch EASYcare Study

We wonder whether the effectiveness of community intervention models for frail elderly people can be enhanced using an alternative way of participant selection. In addition to selecting participants on the basis of age and frailty criteria, we ask the general practitioner to initiate the intervention when a problem requiring action emerges. This problem-based approach may enhance effectiveness because of better timing of the intervention. Others have shown this type of targeting can be effective, albeit in a non-randomised design [24]. General practitioner's and participant's compliance may also benefit, because both have discussed and agreed on the involvement of another health provider. The general practitioner is directly involved in the intervention model which realises more control over the clinical management. Direct involvement of the general practitioner also provides feedback possibilities to better tailor the intervention and it safeguards continuity of care. We presume this continuity prevents the occurrence of negative effects that could result from discontinuation of the intervention. Hypothetically,

the result is that the intensive involvement of health workers than other the general practitioner and regular home care is needed only temporarily.

If an informal caregiver was involved, we actively engaged this person in our intervention. We believe this involvement is a precondition for an effective community intervention model focusing on frail elderly people. However, to our knowledge, this caregiver involvement has not received much attention in the empirical studies of community intervention models.

Objectives

The objective of our study is to determine the effects of the Dutch EASYcare Study Geriatric Intervention Programme (DGIP) compared to regular medical care in improving health-related quality of life in independently living elderly persons and in improving caregiver burden. Moreover, we want to determine the costs of the Dutch EASYcare Study Geriatric Intervention Programme.

Methods/Design

Study design and setting

The study is an observer blinded randomised controlled trial. Pseudo cluster randomisation was used to randomly allocate the participants to the DGIP or to a regular care group. Pseudo cluster randomisation is a randomisation method that aims to prevent both the occurrence of selection bias and contamination in a single design. We will discuss it in more detail below. The Ethical committee of the Radboud University Nijmegen Medical Centre approved of the study.

Study population

54 general practitioners from 36 primary care practices in and around Nijmegen, the Netherlands, were willing to recruit subjects. We started with 38 general practitioners, but increased this number during the recruitment period because of disappointing inclusion rates. During the inclusion period of 21 months 155 eligible participants were randomised. We decided not to include in follow-up and analysis those participants who experienced severe intercurrent disease necessitating hospital admittance, were admitted to a nursing home, died, or withdrew informed consent within one week after randomisation. The possibility of the study to have effect within one week after randomisation was judged as negligible, because it took about a week before nurses started the intervention, and the follow-up visits were judged to be too strenuous for these seriously ill patients. Therefore 151 participants were included in follow-up and analysis; 85 were included in the group that received the intervention model, and 66 were included in the regular care group.

Eligibility criteria

Subjects had to be eligible for participation in our intervention model (table 1). All participants had to be living in their own home or in a home for the aged and had to be 70 years or older.

When we started recruiting participants we applied an age criterion of 75 years or older. Unfortunately, seven months after the start of the recruitment the inclusion rates fell

short of expectations. We decided we were able to broaden the age criterion, because the combination of frailty criteria and a problem driven approach safeguarded selection of eligible participants.

table 1 *Eligibility criteria for Dutch EASYcare Study*

Inclusion criteria
• 70 years of age and over
• The patient lives independently or in a home for the aged
• The patient has a health problem that was recently presented to the general practitioner by the patient or informal caregiver
• The request for help is related to the following problem fields: cognitive disorders, behavioural and psychological symptoms of dementia, mood disorders, mobility disorders and falling, or malnutrition
• The patient/informal caregiver and general practitioner have determined a goal they want to achieve
• Fulfil one or more of these criteria: MMSE (Mini Mental State Examination) equal to or less than 26, GARS (Groningen Activity Restriction Scale) equal to or greater than 25 or MOS-20/subscale mental health equal to or less than 75
Exclusion criteria
• The problem or request for help has an acute nature, urging for action (medical or otherwise) within less than one week
• The problem or request for help is merely a medical diagnostic issue, urging for action only physicians (general practitioner or specialist) can offer
• MMSE < 20 or proven moderate to severe dementia (Clinical Dementia Rating scale [CDR] > 1, 0) and no informal caregiver (no informal caregiver is defined as: no informal caregiver who meets the patient for at least once a week on average)
• The patient receives other forms of intermediate care or health care from a social worker or community-based geriatrician
• The patient is already on the waiting list for a nursing home because of the problem the patient is presented with in our study
• Life expectancy < 6 months because of terminal illness

We restricted participant inclusion to those who scored below maximum (indicating good performance) on at least one of the following instruments: Mini Mental State Examination (MMSE), MOS-20 subscale mental health, or Groningen Activity Restriction Scale (GARS) [25-27]. For the MMSE the cut-off was equal to or less than 26 out of 30, for MOS-20 mental health equal to or less than 75 out of 100, and for GARS the cut-off was equal to or greater than 25. The GARS score ranges 18 to 54, where 18 indicates best functional performance. We excluded participants with an MMSE of less than 20 or a proved moderate to severe dementia (Clinical Dementia Rating scale [CDR] > 1, 0) and no informal caregiver, because we expected serious problems in the acquisition of research data in these persons.

Persons already receiving forms of intermediate care or health care from a social worker or community-based geriatrician were also excluded, because this made it difficult to establish which effect was measured. Receiving home care, however, was not an exclusion criterion.

Persons already on the waiting list for a nursing home, or who had a life expectancy of less than six months, because of terminal illness, were excluded as well.

As a result of a mistake, in one case the age criterion was violated. However, the intervention team agreed that this younger case (age of this participant was 69 years) fitted well into the model. As exclusion was judged to be in disagreement with the ethical treatment of participant data, this participant was kept in follow-up and analysis.

Treatment arms and randomisation

Participants were randomly allocated over two treatments arms: DGIP and regular care. No restrictions were imposed on the care participants were allowed to receive in the regular care group.

Given the nature of our intervention we considered the use of two different allocation procedures available in literature: cluster randomisation or individual randomisation [28]. The use of a cluster randomised design may have had an advantage over the use of an individual randomised design, because of the possible occurrence of contamination in our trial when individual randomisation was applied [29]. On the other hand a cluster randomised design had several disadvantages. The general practitioner would have known the allocation outcome for his cluster after the first patient in a fully cluster randomised design. This might have caused selection bias resulting in incomparability of treatment arms [30, 31]. At the same time we presumed it likely that the recruitment of subjects in the control clusters would progress slowly. Why should a general practitioner bother to refer a patient to a study, when the general practitioner knows already that the patient will enter the control group? There is also evidence for differential recruitment rates in cluster randomisation [32].

We therefore choose to use an innovative two-step pseudo cluster randomisation procedure [28, 33]. First the general practitioners were randomised into two groups; group I and group C. The results of this randomisation were not revealed. Then within each of these groups randomisation at the patient level was carried out. This randomisation was stratified by general practitioner and performed in such a way that in group I the majority (approximately 80%) of the participants received the intervention treatment, while the others received standard treatment. In group C the dysbalance was reversed: the majority received standard treatment and the others got the intervention treatment.

This approach had important advantages. The general practitioner did not know in advance which treatment a patient was going to get, so this reduced the chance of selection bias. It also prevented the occurrence of negative recruitment effects that might have resulted from being randomised to a control cluster. Had the general practitioners known in advance the group they were assigned to (I or C), the predictability of an individual randomisation decision had been larger than in an individually randomised trial. However, the randomisation of general practitioners occurred blinded. In such

a situation, the general practitioner can only gain knowledge on the randomisation proportion through the recruitment of participants. As the number of enrolled patients per practice was expected to be no more than 10, the chances to correctly guess the odds for each individual treatment are limited.

We expect the contamination due to the intervention treatment to be negligible in group C, because there are only a limited number of participants in this group on the experimental treatment. As the majority of the patients is on intervention treatment, the contamination may be a problem in patients in group I who are on standard treatment, but then it probably affects only a small portion of the patients.

A randomisation procedure with adaptive weights (minimisation) was used to ensure a balanced distribution of high versus low percentage of elderly per primary care practice and of the availability of a nurse practitioner in primary care practice in the two groups I and C [34]. The patients were randomised with adaptive weights to get evenly distributed numbers of sex, and presented health problem. A person not related to the study conduct performed the randomisation.

Intervention model: DGIP

General practitioners referred independently living older patients to our model when there was a problem in cognition, nutrition, behaviour, mood, or mobility. The problem had to urge for nursing assessment, coordination of care, or therapeutic monitoring and case management. Requests were rejected if they had an acute nature or if they were purely medical diagnostic requests.

A suitable case is for example a widow living on her own in a flat on the second floor with no elevator. The general practitioner has doubts about her cognitive abilities and she has depressive symptoms as well. This seems to affect her daily functioning, although to what extent is unclear. She has only a daughter to look after her.

After negotiating a preliminary goal with the patient, the referring general practitioner contacted the geriatrician involved in the study. Within two weeks a geriatric specialist nurse visited the patient at home. The instrument EASYcare was applied during this first visit [35]. EASYcare is an instrument for geriatric assessment that consists of items about (instrumental) activities of daily life, cognition, mood, and ends with a goal setting item. The goal initially negotiated by patient and general practitioner was further elaborated in an operational objective. If an informal caregiver was present, the nurse provided this person a caregiver burden assessment and the results were implemented in the care plan.

During maximum three months up to five follow-up visits for additional geriatric evaluation and management were planned. The nurse, geriatrician, and general practitioner frequently discussed the necessary nursing interventions, the effect of the interventions, the level of care that was needed, and the possibilities for reversibility. If necessary the nurse consulted and advised other involved health care workers, such as home care or physical therapist.

We had two nurses and two geriatricians available for the execution of our intervention. We developed guidelines based on best nursing practice for each health problem to structure activities, because literature has pointed at the possibility that the effects of home visiting programmes are related to the home visitor's performance in conducting the visits [36]. Therefore, we structured the intervention in order to diminish this effect, without harming the flexibility of the model. Our guidelines divided the nursing process into four phases: nursing diagnosis, definition of expected outcomes, nursing interventions and assessment of outcomes. Secondly, the guidelines used standardised NANDA (North American Nursing Diagnosis Association), NOC (Nursing outcomes classification) and NIC (Nursing interventions classification) terminology for nursing diagnosis, nursing outcomes and nursing interventions respectively [37-39].

We piloted our intervention model in a feasibility study [40]. With some minor changes, this model was judged to be applicable in the current study.

Data collection and outcome measures

Within one week after referral a researcher (RM, ME) interviewed patients at home to obtain written informed consent and to collect baseline demographic characteristics and data on general health conditions. If the participant was not able to give informed consent we asked a proxy to do so. The participants always gave verbal assent and did not reject the measurements. Before the interview the participant received a written confirmation of the appointment and a questionnaire. We asked the participant to fill out the questionnaire before the appointment. If the participant was unable to fill out the questionnaire independently, we allowed help from another person. In some cases the interviewer filled out the questionnaire during the interview. We recorded the amount of help the participant received in filling out the questionnaire.

The participants provided data on the following measures: age, gender, type of residence, and the use of home care. Also, data were collected on functional abilities, cognitive condition, mobility, health-related quality of life, and loneliness.

If an informal caregiver was available we collected data on informal caregiver characteristics using a questionnaire. We collected data on type and amount of care provided, time spent on caring, and caregiver burden.

These measurements are repeated three and six months after inclusion. The same researcher that performed the baseline visit carries out these interviews. This researcher is not involved in the intervention nor does the researcher know the allocation decision. After each follow-up visit the researcher indicates whether blinding remained intact or not.

Primary outcome measures relating to participant characteristics are functional performance in (independent) activities of daily living measured using Groningen Activity Restriction Scale and mental health using subscale mental health MOS-20. Primary outcome measure in informal caregivers is caregiver burden using the Zarit Burden Interview (ZBI) [41]. An overview of secondary outcomes and a complete list of all measurements are provided in table 2.

Process evaluation

We collect data on the following set of process variables: the content of the intervention programme, the adherence of participants and informal caregivers in the intervention group to advices given during an intervention, experiences of participants and informal caregivers with the intervention model, and data on general practitioner care and care of other involved professionals in both treatment arms.

table 2 *Outcome measures*

Variable	Primary outcome	Secondary outcome	Background variable	Instrument	Measured at	T_0	T_1	T_2
Functional performance (ADL/ IADL)	☑	☐	☐	GARS-3 [27]		☑	☑	☑
• Mobility	☐	☑	☐	Timed up and go test [44]		☑	☑	☑
Health-related quality of life	☐	☑	☐	MOS-20 [26]		☑	☑	☑
Mood	☑	☐	☐	Subscale mental health MOS-20		☑	☑	☑
Well-being	☐	☑	☐	Cantril self-anchoring ladder [45]		☑	☑	☑
				Dementia Quality of Life questionnaire[46]		☑	☑	☑
				question general life satisfaction		☑	☑	☑
Cognition	☐	☑	☐	MMSE [25]		☑	☐	☑
Social functioning	☐	☑	☐	Loneliness scale de Jong-Gierveld [47]		☑	☑	☑
Mortality	☐	☑	☐			☐	☑	☑
Housing conditions/sort of residence	☐	☑	☐	Own questionnaire		☑	☑	☑
Subjective treatment effects (participant, informal caregiver)	☐	☑	☐	Patient Enablement Instrument [48]		☐	☑	☐
Burden informal caregiver	☑	☐	☐	Zarit Burden Interview [41]		☑	☑	☑
	☐	☑	☐	Questions taken from "Zorgkompas Mantelzorger" [49]		☑	☑	☑
Time spend on care (informal caregiver)	☐	☑	☐	Own questionnaire		☑	☑	☑
Age (participant, informal caregiver)	☐	☐	☑	Own questionnaire		☑	☐	☐
Sex (participant, informal caregiver)	☐	☐	☑	Own questionnaire		☑	☐	☐
Socio-economic status	☐	☐	☑	Own questionnaire, classify using ISEI-92 [50]		☑	☐	☐
• (Former) occupation	☐	☐	☑	Own questionnaire, classify using SBC-92 [50]		☑	☐	☐
Nativity	☐	☐	☑	Own questionnaire		☑	☐	☐
Co-morbidity	☐	☐	☑	Cumulative Illness Rating Scale-Geriatrics (CIRS-G) [51] from medical history in general practitioner Information System		☑	☐	☐
Use of home care	☐	☐	☑	Own questionnaire		☑	☑	☑

T_0 is baseline measurement

T_1 is first follow-up measurement, after 3 months

T_2 is second follow-up measurement, after 6 months

We collect data on the content of the intervention process, because this may help to identify which programme characteristics are most beneficial. An abstract form is used to extract this information from the nursing records after completion of all individual interventions. We extract information on treatment goals, nursing diagnoses (NANDA)[37], nursing interventions (NIC)[38], nursing outcomes (NOC)[39], and the employed diagnostic instruments.

Compliance of participants and informal caregivers is an important determinant of carrying out a successful intervention. When an individual intervention is finished the nurse that executed the intervention indicates in an MS Access® form which of a number of pre-specified advices were given. Another nurse calls the participant or informal caregiver one month later to check compliance on these advices.

We score subjective treatment effects in treatment group using a questionnaire that participants and informal caregivers filled out after the first follow-up visit.

Data on general practitioner care will be collected in both treatment arms from the information that is routinely available from the General Practice's Information System (Huisartsen Informatie Systeem). We collect the following data: medical history using ICPC-2 (International Classification of Primary Care) [42], number and content of contacts during six months of follow-up using ICPC-2, number and nature of referrals, and medication using ATC classification (Anatomical Therapeutic Chemical drug classification) [43]. Data on the use of home care are collected in the participant questionnaire. The data on general practitioner care will be collected at the end of the follow-up period. These data are collected in order to be able to clarify the observed intervention effect and to establish costs.

Costs

To be able to calculate costs, data will be collected on the following cost variables. Nurses will register the time spent on the intervention using the MS Outlook® agenda. They will register the number of visits per participant. They also register the time spent on consultation, phone calls, travelling, and administration.

Data on the workload of the general practitioner and the geriatrician will be extrapolated from the workload of the nurses. The data we collect on the care provided were already described in the paragraph "process evaluation". Finally, we will derive salary costs, administrative costs, and costs for materials.

Sample size considerations

A change in the primary outcome measure of functional performance (GARS-3) of 4.5 points on a scale ranging from 18 (complete independence) to 54 (complete dependence) can be found with a power (1-β) of 0.80 and α (two sided) of 0.05 in comparing two groups of 77 subjects, when pseudo cluster randomisation is applied. We use a standard deviation of 8.5, which we calculated from a pilot study. This standard deviation is well in the range of the measures of spread other studies have found [27]. A mean increase of 4.5 points is chosen as clinically relevant, because a 4.5 point increase of the overall score indicates an improvement of 25% of all items by one functional class (each item's score is classified as follows: completely dependent 3 point, partly dependent 2 points

and completely independent 1 point). Cluster size is estimated to be approximately 10 participants per general practitioner. The exact calculations and considerations are extensively described in Teerenstra et al [33].

Statistical analysis

Descriptives will be used to assess comparability of both intervention and control group for background and confounding variables. Our primary analysis will focus on the treatment arms' differences in the primary outcome measures' changes from baseline (GARS, MOS-20 subscale mental health, and Zarit Burden Interview) at three months of follow-up (T_1). This will be done in intention-treat-analysis. We will use mixed linear model analysis (Proc Mixed in SAS® 8) to quantify these differences. We will account for clustering at the level of the general practitioner through the addition of a random intercept for general practitioner to the three models. The baseline measurements of GARS, MOS-20 subscale mental health, and Zarit Burden Interview will be added to the respective models as a covariate. The factors we stratified for in the randomisation (general practitioner characteristics, sex of participant, and participant's presented health problem) will also be added to the models as covariates. No further corrections will be made. A conditional analysis of the treatment arms' differences in changes from baseline at six months (T_2) will be performed if there is a significant effect at T_1. Apart from replacing the scores at three months with those at six months the same three models will be used.

The secondary analyses will be performed on the treatment arms' differences in time trend of the primary outcome measures GARS, MOS-20 subscale mental health, and Zarit Burden Interview during follow-up. Secondary analysis will further focus on the differences between treatment arms of the secondary outcome measures at three and six months of follow-up. Kaplan-Meier estimates and hazard ratios will be used to quantify the intervention's effect on living conditions and mortality. Subgroup analyses will be performed for the following subgroups: living in one's own home versus living in a home for the aged, and higher versus lower levels of cognitive function. All analyses will be performed in SAS® 8.

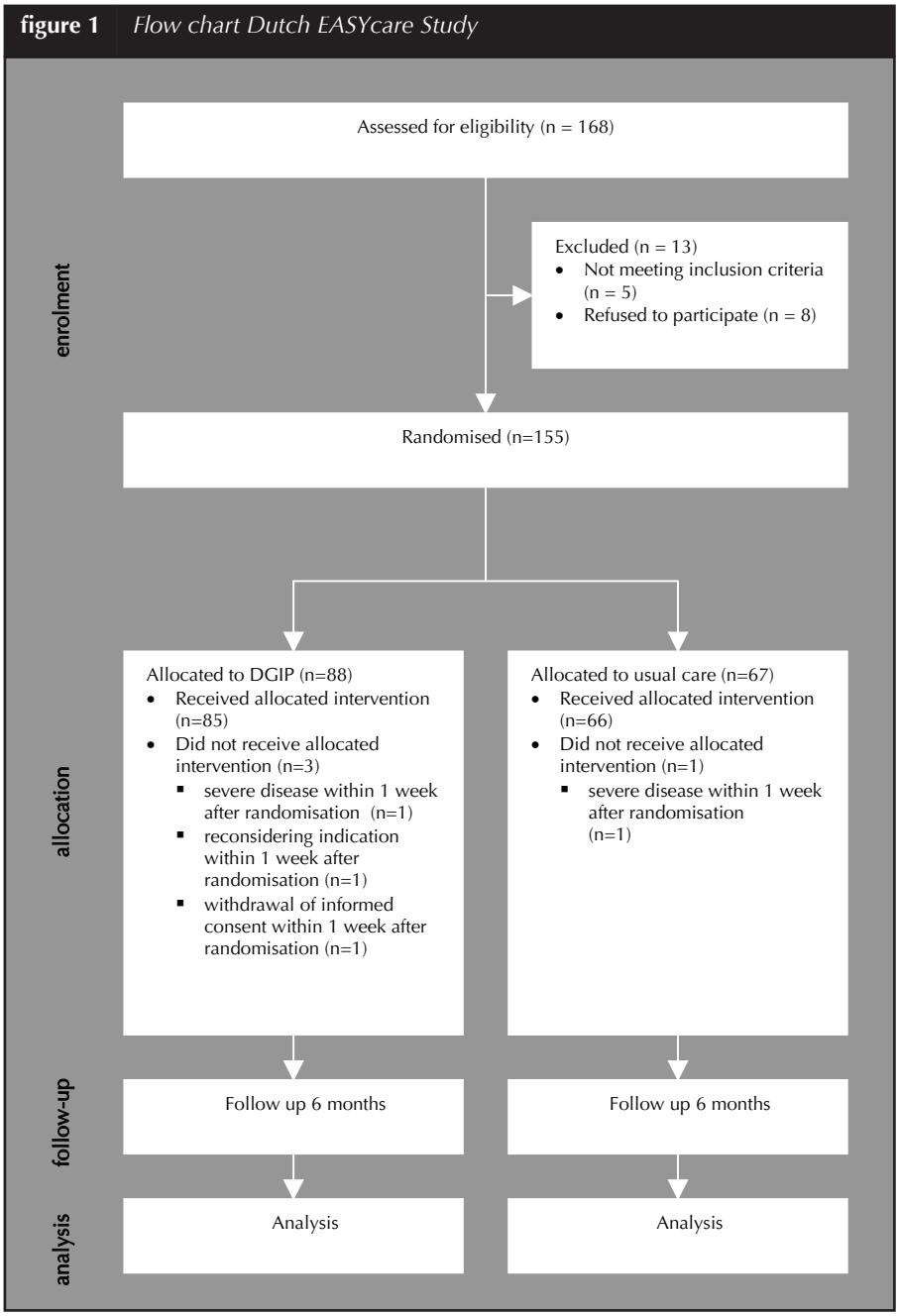
Discussion

In this paper we presented the design of a randomised controlled trial that evaluates the effects of a community intervention model for frail elderly people living on their own. The design of this study has shown to be very challenging.

Although the recruitment of the participants took much effort, we have included a number of subjects that should be large enough to provide reliable answers to our research questions (figure 1).

Our participants were selected using a problem-based approach in which the general practitioners decided in cooperation with the geriatrician which patients were suitable for this intervention model. This participant selection procedure satisfied; only a minor number of the referred patients did not meet our eligibility criteria based on frailty and age. Probably, piloting our intervention model was important to achieve this.

figure 1 *Flow chart Dutch EASYcare Study*



This flow chart summarises the progress through the phases of the Dutch EASYcare Study until the allocation of participants to each treatment arm

As discussed earlier, there is a lack of insight into the content of most community intervention models studied. We decided to use standard terminology such as ICPC, NANDA, NIC, NOC and ATC codes to provide insight into our intervention when used in practice. This makes detailed insight possible using terminology others can understand well.

The selection of the best randomisation method was a final major issue we had to deal with and that took much of our time. We think this randomisation procedure satisfies. Nevertheless, we will closely monitor and report in future papers how the randomisation procedure performs in practice.

References

1. von Faber M, Bootsma-van der WA, van Exel E et al. Successful aging in the oldest old: Who can be characterized as successfully aged? *Arch Intern Med* 2001 December 10;161(22):2694-700.
2. Rowe JW, Kahn RL. Human aging: usual and successful. *Science* 1987 July 10;237(4811):143-9.
3. Rubenstein LZ. An overview of comprehensive geriatric assessment: Rationale, History, Program Models, Basis Components. In: Rubenstein WBMEK, editor. Geriatric Assessment Technology: The State of the Art. 1996. p. 1-312.
4. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993 October 23;342(8878):1032-6.
5. Salvage AV, Jones DA, Vetter NJ. Options of people aged over 75 years on private ad local authority residential care. *Age Ageing* 1989;18:380-6.
6. WHO. Active ageing: a policy framework. Geneva: World Health Organisation; 2002.
7. Melis RJF, Olde Rikkert MG, Parker SG, van Eijken MIJ. What is intermediate care? *BMJ* 2004 August 14;329(7462):360-1.
8. Rubenstein LZ, Stuck AE. Preventive home visits for older people: defining criteria for success. *Age Ageing* 2001 March;30(2):107-9.
9. Carpenter I, Gladman J, Parker S, Potter J. Clinical and research challenges of intermediate care. *Age Ageing* 2002;31:97-100.
10. Black C, Black D, Alberti G. Intermediate care: statement from the Royal College of Physicians of London. London: The Royal College of Physicians; 2000 Oct 5.
11. Parker G, Bhakta P, Katbamna S et al. Best place of care for older people after acute and during subacute illness: a systematic review. *J Health Serv Res Policy* 2000 July;5(3):176-89.
12. Walsh B, Steiner A, Warr J, Sheron L, Pickering R. Nurse-led inpatient care: opening the 'black box'. *Int J Nurs Stud* 2003 March;40(3):307-19.
13. Elkan R, Kendrick D, Dewey M et al. Effectiveness of home based support for older people: systematic review and meta-analysis. *BMJ* 2001 September 29;323(7315):719-25.
14. Shepperd S, Iliffe S. Hospital-at-home versus in-patient hospital care. *Cochrane Database Syst Rev* 2000;(2):CD000356.
15. Parker G, Peet S. Position paper: intermediate care (produced for the Department of Health. 2001.
16. Wieland D. The effectiveness and costs of comprehensive geriatric evaluation and management. *Crit Rev Oncol Hematol* 2003 November;48(2):227-37.
17. Fletcher AE, Price GM, Ng ES et al. Population-based multidimensional assessment of older people in UK general practice: a cluster-randomised factorial trial. *Lancet* 2004 November 6;364(9446):1667-77.
18. Bula CJ, Alessi CA, Aronow HU et al. Community physicians' cooperation with a program of in-home comprehensive geriatric assessment. *J Am Geriatr Soc* 1995 September;43(9):1016-20.
19. Newbury J, Marley J. Preventive home visits to elderly people in the community. Visits are most useful for people aged >/= 75. *BMJ* 2000 August 19;321(7259):512.
20. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. *JAMA* 2002 February 27;287(8):1022-8.
21. van Rossum E, Frederiks CM, Philipsen H, Portengen K, Wiskerke J, Knipschild P. Effects of preventive home visits to elderly people. *BMJ* 1993 July 3;307(6895):27-32.
22. Egger M. Commentary: When, where, and why do preventive home visits work? *BMJ* 2001;323(7315):724-5.
23. Stuck AE, Beck JC, Egger M. Preventing disability in elderly people. *Lancet* 2004 November 6;364(9446):1641-2.
24. Landi F, Gambassi G, Pola R et al. Impact of integrated home care services on hospital use. *J Am Geriatr Soc* 1999 December;47(12):1430-4.
25. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975 November;12(3):189-98.
26. Kempen GI. Het meten van de gezondheidstoestand van ouderen. Een toepassing van een Nederlandse versie van de MOS-schaal. [Assessment of health status of the elderly. Application of a Dutch version of the MOS scale]. *Tijdschr Gerontol Geriatr* 1992 August;23(4):132-40.
27. 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 December;43(11):1601-10.
28. Borm GF, Melis RJF, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. *Stat Med* 2005 December 15;24(23):3535-47.
29. Donner A, Birkett N, Buck C. Randomization by cluster. Sample size requirements and analysis. *Am J Epidemiol* 1981 December;114(6):906-14.
30. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001 February 10;322(7282):355-7.
31. Jordhoy MS, Fayers PM, Ahlner-Elmqvist M, Kaasa S. Lack of concealment may lead to selection bias in cluster randomized trials of palliative care. *Palliat Med* 2002 January;16(1):43-9.
32. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ* 2003 October 4;327(7418):785-9.
33. Teerenstra S, Melis RJF, Peer PG, Borm GF. Pseudo cluster randomization dealt with selection bias and contamination in clinical trials. *J Clin Epidemiol* 2006 April;59(4):381-6.
34. Altman DG, Bland JM. Treatment allocation by minimisation. *BMJ* 2005 April 9;330(7495):843.
35. Richardson J. The Easy-Care assessment system and its appropriateness for older people. *Nurs Older People* 2001 October;13(7):17-9.
36. Stuck AE, Minder CE, Peter-Wuest I et al. A randomized trial of in-home visits for disability prevention in community-dwelling older people at low and high risk for nursing home admission. *Arch Intern Med* 2000 April 10;160(7):977-86.
37. *NANDA Nursing diagnoses: definitions and classification*. 6 ed. Philadelphia: NANDA; 2005.
38. *Nursing Interventions Classification (NIC)*. 3 ed. St Louis: Mosby; 2000.
39. *Nursing Outcomes Classification (NOC)*. 3 ed. St Louis: Mosby; 2003.
40. van Eijken MIJ, Melis RJF, Wensing M, Olde Rikkert MGM, van Achterberg T. Contents of a GP supportive geriatric intermediate care programme: a descriptive study. *submitted* 2005.
41. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist* 1980 December;20(6):649-55.
42. *ICPC. International Classification of Primary Care*. Oxford: Oxford University Press; 1987.
43. About the ATC/DDD system. 2005. Oslo, WHO Collaborating Centre for Drug Statistics Methodology.
44. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 February;39(2):142-8.
45. Cantril H. *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press; 1965.
46. Brod M, Stewart AL, Sands L, Walton P. Conceptualization and measurement of quality of life in dementia: the dementia quality of life instrument (DQoL). *Gerontologist* 1999 February;39(1):25-35.
47. de Jong Gierveld J, Kamphuis FH. The development of a Rasch-type loneliness scale. *Appl Psychol Measurement* 1985;9:289-99.
48. Howie JG, Heaney DJ, Maxwell M, Walker JJ. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. *Fam Pract* 1998 April;15(2):165-71.
49. Duijnstee M, Guldemond H, Hendriks L. *Zorgkompas voor mantelzorgers van ouderen en chronisch zieken. Leidraad voor het in kaart brengen van de belasting van zorgende familieleden [care compass for informal carers of elderly people and people with a chronic illness. A guide for mapping the burden of caring family members]*. 1 ed. Utrecht: Nederlands Instituut voor Zorg en Welzijn/NIZW; 2001.
50. Bakker B, Sieben I, Nieuwbeerta P, Ganzeboom H. Maten voor prestige, sociaal-economische status en sociale klasse voor de standaard beroepen classificatie 1992 [Scales for prestige, socio-economic status and social class for the Standard Occupational Classification 1992]. *Sociale Wetenschappen* 1997;40(1):1-22.
51. Miller MD, Towers A. *A Manual of Guidelines for Scoring the Cumulative Illness Rating Scale for Geriatrics (CIRS-G)*. Pittsburg: University of Pittsburgh; 1991.

Chapter 4

Pseudo cluster randomisation dealt with selection bias and contamination in clinical trials

S. Teerenstra, R.J.F. Melis, P.G.M. Peer, G.F. Borm

J Clin Epidemiol 2006;59:381-6

Objective: When contamination is present, randomisation on a patient level leads to dilution of the treatment effect. The usual solution is to randomise on a cluster level, but at the cost of efficiency and more importantly, this may introduce selection bias. Furthermore, it may slow down recruitment in the clusters that are randomised to the "less interesting" treatment. We discuss an alternative randomisation procedure to approach these problems.
Study design and setting: Pseudo cluster randomisation is a two-stage randomisation procedure that balances between individual randomisation and cluster randomisation. For common scenarios, the design factors needed to calculate the appropriate sample size are tabulated.
Results: A pseudo cluster randomised design can reduce selection bias and contamination, while maintaining good efficiency and possibly improving enrolment. To make a well-informed choice of randomisation procedure, we discuss the advantages of each method and provide a decision flow chart.
Conclusion: When contamination is thought to be substantial in an individually randomised setting and a cluster randomised design would suffer from selection bias and/or slow recruitment, pseudo cluster randomisation can be considered.

Introduction

The golden standard for randomisation is assigning individuals at random to experimental conditions (individual randomisation). This maximises the probability that (un)known confounding variables will be evenly distributed over treatment groups. However, when such an allocation procedure is applied, the subjects on one treatment may come into contact with elements of the other treatment and vice versa. Consequently, we may not observe the pure treatment effect, but a contaminated one: some elements of one treatment mix into the other treatment and vice versa. If such contamination is thought to be strong enough to pull the observed treatment difference below statistical significance, there are two options. Firstly, the panacea of increasing the sample size [1]. This can endow the contaminated effect with statistical significance, but it cannot prevent the contamination from pulling the observed effect below *clinical* relevance.

Secondly, contamination can be avoided altogether by using cluster randomisation. Subjects who are likely to contaminate each other are grouped into clusters and entire clusters are randomised to one of the treatments. However, this clustering decreases the effective sample size due to the positive correlation between the subjects within the clusters.

Two caveats are attached to cluster randomisation. Often, it is impossible to recruit all the subjects before the clusters are randomised. As a result, it is known in advance which treatment the patients in a certain cluster will receive and this may lead to selection bias [2]. The other problem is the risk of slow or stagnating patient recruitment in the clusters that are going to receive what the patients or recruiters consider to be the "less interesting" treatment.

To avoid the above-described objections to cluster randomisation, Borm et al. [3] devised a randomisation procedure that balances between individual randomisation and cluster randomisation.

In the two sections below, we describe the method of pseudo cluster randomisation, while in section 4, we present a formula for sample size calculation. In the discussion, we compare the merits and demerits of pseudo cluster randomisation to those of ordinary cluster randomisation and individual randomisation and provide guidelines to make a well-informed choice between them.

Pseudo cluster randomisation

Pseudo cluster randomisation is a two-stage randomisation procedure that combines randomisation on a cluster level with randomisation on a patient level. Suppose that a trial is set up to evaluate two treatments, *T* and *S*. In the first step, the clusters are randomised into two groups: half of the clusters go into cluster group *t* and the other half go into cluster group *s*. The allocation of each cluster is not revealed. In the second step, subjects are allocated to treatment *T* or *S* in the following way: within each cluster in cluster group *t*, subjects are randomised in such a way that the majority, say a fraction *f* x 100%, receive treatment *T* and the remainder, (1-*f*) x 100%, receive treatment *S*. In cluster group *s*, the fractions are reversed, i.e. the majority of each cluster receive treatment *S* and the remainder receive treatment *T*.

The Dutch EASYcare trial

Pseudo cluster randomisation is being employed in the Dutch EASYcare trial to address the problems associated with individual randomisation and cluster randomisation [4]. This is an observer-blinded trial that aims to compare an intermediate care programme for elderly people with geriatric problems to the traditional care as provided by general practitioners. General practitioners recruited elderly patients with a geriatric problem for the study. The intermediate care programme is performed by a specialist geriatric nurse and consists of up to six home visits for geriatric evaluation and management. general practitioner participation in this treatment arm is limited to regular case evaluations with the nurses.

In a design with individual randomisation, the general practitioner would have as many patients on the new treatment as on the standard treatment. Contamination may then arise on a general practitioner level: due to feedback from the nurses about the elements and effects of the intermediate care programme, the general practitioner may gradually acquire knowledge and skills to implement (parts of) the intermediate care programme in the care of other patients, including those originally allocated to traditional care. On the other hand, contamination on a patient level is unlikely, because the patients are not likely to meet each other.

The above-described contamination can be avoided if each general practitioner is concerned with one treatment only (cluster randomisation). However, in this trial, as is often the case with cluster randomised designs, it was not possible to recruit all the patients before the clusters were randomised. The treatment of a geriatric problem could not be postponed until after completion of recruitment. Therefore, (after the first patient) the general practitioner would know which treatment all the following patients would be assigned to and this might influence the general practitioner's choice of patients (selection bias). Also, as the main motivation for general practitioners to take part in this trial was their interest in the new treatment, the general practitioners in the control clusters may be less motivated to recruit patients and they might produce data of poorer quality [5]. Thus, recruitment may be poor in the control clusters, which can jeopardise the power of the study [6].

We decided to use pseudo cluster randomisation to deal with this dilemma. First the general practitioners were randomised into two groups on a 1:1 basis: *t* (treatment) or *s* (standard). The results of this randomisation were not revealed. Then, within clusters of group *t* (or *s*), patients were randomised to the new treatment or standard care at a ratio of 0.8 : 0.2 (0.2 : 0.8, respectively). One of the merits of this procedure compared to cluster randomisation was that none of the general practitioners ended up with only patients

on the uninteresting (standard) treatment, which may have improved recruitment. In addition, the general practitioner did not know in advance which treatment the next patient would receive, which tempered selection bias. As recruitment proceeded, the general practitioner might have noticed the imbalance in treatment ratio and tried to guess the chance of either treatment, but predictability would always be lower than in cluster randomisation. Predictability was further reduced in this trial by keeping the clusters small. Thus, it was harder to guess the treatment ratio before the end of the recruitment period.

In comparison with individual randomisation, pseudo cluster randomisation may lead to less contamination in the Dutch EASYcare Study. At the start of the study there was no contamination, because the general practitioners did not have the knowledge or skills to apply (elements of) the invention. During the study, the general practitioner does not execute the intervention and is only provided with information about the treatment through the evaluations with the nurse. As the EASYcare intervention is a multi-faceted, tailored treatment, the general practitioner must have had feedback concerning several intermediate care patients before he or she has any grasp of the elements of the intervention. Moreover, not only passive knowledge will have to be acquired, but also new skills and attitudes will have to be learned in order to copy the intervention successfully [7]. Therefore, it is unlikely that having 2 patients allocated to intermediate care within a general practice would cause any substantial contamination, but having 5-8 patients allocated to intermediate care could be a concern. In other words, there will be hardly any contamination of the results of the patients on the majority treatment within a cluster, whereas there may be substantial contamination in the patients on the minority treatment. The latter contamination, however, would only affect a relatively small number of the participants. Therefore, we considered pseudo cluster randomisation to be an appropriate means to reduce contamination.

Sample size calculations

Sample size calculations for pseudo cluster randomised designs follow the same procedure as those for ordinary cluster randomised designs.

First, calculate the sample size *N* with the usual formula for an individually randomised trial:

$$(1) \qquad N = 2(z_{\alpha/2} + z_{\beta})^2 \sigma^2 / \delta_D^2$$

where $z_{\alpha/2}$ and z_{β} are the quantiles of the normal distribution, given that the type I error rate is α (two-sided) and the type II error rate is β. In addition, σ is the standard deviation and δ_D is the treatment difference expected in the chosen design.

Second, multiply the above sample size by the appropriate design factor, which not only depends on the intracluster correlation ρ, but also on the majority fraction *f* and cluster size *n*. For *f*=0.8, the design factors of pseudo cluster randomisation are tabulated in Table 1.

table 1	*Design factors based on the minimal variance estimator in Borm et al[3] by cluster size n and by intracluster correlation ρ*						
	Intracluster correlation ρ						
Cluster size n	0.005	0.01	0.05	0.1	0.2	0.3	0.4
10	1.01	1.02	1.08	1.11	1.08	0.99	0.87
20	1.03	1.05	1.17	1.20	1.14	1.03	0.90
30	1.04	1.08	1.22	1.24	1.17	1.05	0.91
40	1.06	1.10	1.26	1.27	1.19	1.06	0.92
50	1.07	1.13	1.29	1.30	1.20	1.07	0.92
80	1.11	1.18	1.34	1.33	1.22	1.08	0.93
100	1.13	1.21	1.36	1.34	1.22	1.08	0.93

For other values of *f*, *n* and ρ, the following formula can be used:

$$(2) \qquad \text{Design Factor}_{pc} = \frac{1+(n-1)\rho}{1+4f(1-f)n\rho(1-\rho)}$$

which is derived from the minimal variance estimator for pseudo cluster randomisation in [3]. The total sample size for the trial will be $2\times N\times$ Design Factor$_{pc}$. Half of this sample size will be allocated to treatment *T* and the same number to treatment *S*. The total number of subjects in each cluster group (i.e. *t* or *s*) is $N\times$ Design Factor$_{pc}$ by design, while both cluster groups contain $N\times$ Design Factor$_{pc}/n$ clusters.

The analysis may be performed with methods used for cluster randomised trials [8-10].

Impact of the choice of randomisation procedure on sample size

The Dutch EASYcare study [4] was powered on GARS-3, a functional performance measure for daily life activities that ranges from 18 to 54 [11]. A pilot study on the target population showed that the GARS had a standard deviation of σ=8.5 (and a mean μ=35). It was expected that the intermediate care programme would improve a patient's capacity to perform daily life activities and result in a higher average score on the GARS that could amount to δ=4.5 in the absence of contamination. At a two-sided significance level of 5% (α=0.05), the objective was to achieve a power of 80% (β=0.2). Below, we describe the impact of each randomisation scenario on the total sample size.

Impact of cluster randomisation

In a cluster randomised setting, contamination is absent and the expected treatment difference is δ_D=4.5. However, clustering introduces a positive correlation between the patients in a cluster. This is expressed by the intracluster correlation coefficient, which was anticipated to be ρ=0.05 in this trial. Following the usual procedure, the sample size of each treatment arm is calculated using formula (1), which yields 56. This value is then corrected by multiplying it with the design factor 1+(*n*-1) ρ (see [8]). As the maximum cluster size expected was *n*= 10, the result was a total sample size of 164.

Impact of individual randomisation

In an individually randomised design, contamination is substantial and it is anticipated to dilute the expected treatment difference to δ_D=3.5. However, no clustering takes

place and using (1), the sample size of each treatment arm is 93, i.e. the total sample size is 186.

Impact of pseudo cluster randomisation

Pseudo cluster randomisation was used in the EASYcare trial. Thus, the rate of contamination was expected to decrease compared to the individually randomised situation, although not as far as the uncontamined level (δ_D=4.5) of a cluster randomised setting. An expected treatment difference of δ_D=4 was deemed realistic. Following the procedure outlined in section 4, the sample size of each treatment arm was calculated to be 71 using (1).

After correction with the appropriate design factor 1.08 in Table 1 (n=10, ρ=0.05, f=0.8), the total sample size was 154.

The above calculations showed that pseudo cluster randomisation had an additional advantage in the Dutch EASYcare trial. The sample size was smaller than that of cluster randomisation and individual randomisation. This is not uncommon as shown by the calculations in [3].

Discussion

It is necessary to carefully weight selection bias, contamination and practical considerations in the decision whether to use individual, pseudo cluster or cluster randomisation.

Choice of randomisation method

In some cases, there are compelling reasons to use a particular randomisation procedure. For instance, if a few subjects in a cluster on one treatment are already capable of contaminating all the other subjects on the other treatment, then it may be wiser to rely on cluster randomisation. Other examples in which cluster randomisation may be mandatory occur when new working protocols or nursing methods are introduced on hospital wards, as these cannot be split up. Also, cluster randomisation is sometimes needed to ensure compliance with protocols, for example when pseudo cluster or individual randomisation involves such high (administrative) workload and strict discipline that compliance becomes questionable. Finally, the lower expense of administering only one treatment condition at all sites may encourage the use of a cluster randomised design.

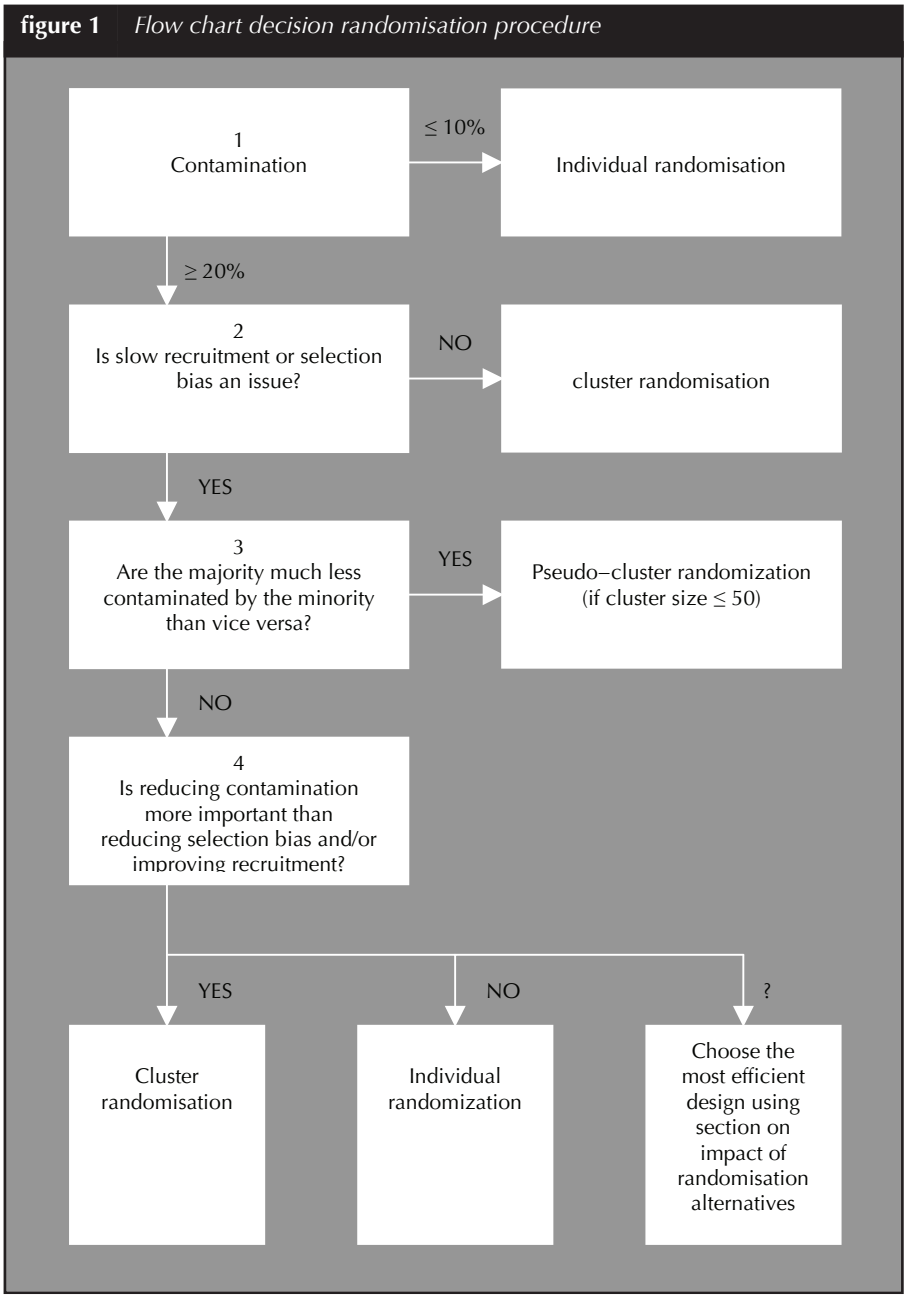
In cases in which both individual and cluster randomisation are feasible options, we present some guidelines that may help to choose between the different types of randomisation (Figure 1).

Consider the comparison of two treatments that have an uncontaminated difference δ in the outcome of interest.

(1) Suppose that in individual randomisation, contamination occurs that dilutes the treatment difference by a factor C, i.e. the treatment difference expected in the individually randomised situation drops to $\delta(1-C)$. Slymen *et al.* [12] showed that the sample size has to be increased by factor $1/(1-C)^2$ to compensate for the loss of power due to this contamination. Thus, if the contamination rate is C=10% the sample size has to be increased by 23%; for $C=20\%$ the increase would be 56%. Therefore, small contamination rates of up to 10% ($C=0.1$) may be compensated for by increasing the sample size by factor $1/(1-C)^2$. If contamination rates are substantial (20% or more), the sample size also has to be increased substantially and (pseudo) cluster alternatives become attractive.

(2) When contamination is large, but recruitment in the "less interesting" clusters or

figure 1 *Flow chart decision randomisation procedure*



selection bias is not an issue, ordinary cluster randomisation is the preferred method. (3) When recruitment and selection bias pose a problem, however, pseudo cluster randomisation may be the recommended choice, especially when contamination of the majority in a cluster (on the one treatment) by a minority in the same cluster (on the other treatment) is considerably smaller than vice versa.

The size of the cluster may play a role in improving recruitment speed and reducing selection bias and contamination. Consider, for instance, the situation in the Dutch

EASYcare trial. As recruitment progresses, the general practitioner might be able to guess with increasing precision how big the chance is of a particular treatment in his or her cluster. At the point where the general practitioner is convinced that most of the patients are on one of the treatments, he or she may still become demotivated to recruit if this is the treatment that he or she finds uninteresting. An even worse effect of this conviction is that it may influence the general practitioner's choice of patients, which may reinduce selection bias. Furthermore, if the cluster size increases, the number of patients on the minority treatment also increases and eventually this may still cause substantial contamination of the majority by the minority. Therefore, pseudo cluster randomisation seems to be feasible for moderate cluster sizes of say between 6 and 50 (which also depends on the enrolment speed). Another good reason not to use pseudo cluster randomisation for large clusters is that individual randomisation is more efficient than pseudo cluster randomisation for large clusters, even though the contamination rate in individual randomisation is larger [3].

(4) If pseudo cluster randomisation is not expected to reduce contamination substantially, three options are available. If reducing contamination is more important, then cluster randomisation is the preferred procedure. In the case of slow recruitment or selection bias, the best choice would be individual randomisation. In cases in which it is unclear what is the most important, efficiency may be the main lead. As in section 4.1, the sample sizes for the different randomisation scenarios can then be worked out and compared.

Choice of *f*

There are two counterbalancing requirements for *f*. On the one hand, the imbalance between the two treatments has to be relatively small (i.e. *f* is close to 0.5) to avoid the selection bias that might arise because of predictability. On the other hand, in order for pseudo cluster randomisation to reduce contamination, the treatment given to the majority of patients in each cluster must be (much) less contaminated by the presence of a minority on the other treatment than vice versa. The latter means that *f* must be closer to 1. Therefore, the value chosen for *f* should be as small as possible as long as the above assumption can still be sustained. In most cases, f=0.8 seems to be a good choice [3].

Reduction of contamination

To reduce contamination by means of pseudo cluster randomisation, the condition mentioned in question box 3 (flow chart 1) needs to be fulfilled i.e. if only a few people in a cluster receive a certain treatment, the outcome of the much larger group who receive the other treatment is much less affected than vice versa. Whether this assumption is reasonable in a specific trial, has to be judged on a case-by-case basis. However, some general remarks can be made.

When contamination is absent at the start of the study, contamination can only result from exposure to the "other" treatment. The relationship between exposure and contamination will follow a certain response or learning curve whose shape depends on the nature of the treatment and exposure. For example, simple information, interventions, and advices are easily transferred verbally or visually which suggests that slight or a single exposure might already result in severe contamination. On the other hand, complex multi-dimensional treatments have a gradual learning curve, as several stages have to be gone through: from acquiring knowledge, via competence and performance to actual successful application of the complex treatment (cf. Miller's pyramid [7] and Reynolds's learning model [13, 14]). Therefore, limiting the driving force of this learning process, which

in our context is the exposure, may limit the contamination effect.
Another aspect is the level at which the contamination arises: this may be the subject level (e.g. patients) or the cluster level (e.g. general practitioners).
In the first case, if there is intense contact between the patients within a cluster (such as within families), contamination by a small proportion of the patients may be considerable and pseudo cluster randomisation would not reduce contamination. In situations with less intensive contacts, pseudo cluster randomisation may be useful. An example is the evaluation of a training programme for the prevention of falls in people living in homes for the elderly. Participants enter the study after their first fall and the treatment may consist of assessment by a physiotherapist who sets up a training programme and advice from an ergotherapist about removing obstacles in the living area and using walking aids, e.g. a walker or a rollator. If only a few people in a home take part in the training, this may not affect the others, so contamination of the large subset may be negligible. Vice versa, if only a few people in another home do not have the training, this may not influence the results of the majority who do take the training.
The Dutch EASYcare trial (section 3) is an example of a trial in which contamination resides on the cluster level. In this trial, general practitioners do not provide the geriatric intermediate care treatment, but may learn the method in the process of several feedback meetings. The more patients who are randomised to the intermediate care at their practice, the more likely these general practitioners are to reach the stage of mastering the intervention and applying it to other patients, which results in contamination.
More trials such as these are likely to be conducted in view of the current trend in care towards multi-disciplinary, individually tailored treatments [15]. These treatments often deal with (multiple) chronic diseases, multiple co-morbidities and often aim to optimally use the capabilities that are still available. Another example is coaching after heart failure by a specialist cardiology nurse, physiotherapist, dietician and social worker in addition to the usual contact with a cardiologist and/or the patient's own general practitioner. The standard care in these contexts is usually provided by mono-disciplinary specialists or all-round medical doctors (general practitioners at practices, nursing homes or homes for the elderly), who use mostly specific, complaint-centred ad-hoc treatments. Although these doctors do not execute the complex interventions themselves, they still supervise them and therefore receive feedback. Consequently, they may unintentionally learn and then use the method on other patients, which results in contamination. In such cases, pseudo cluster randomisation may be well worth considering.

Conclusion

When the risk of selection bias or slow recruitment in a cluster randomised setting competes against contamination arising in individual randomisation, pseudo cluster randomisation may be a good alternative. It can simultaneously minimise both selection bias and contamination and help to improve enrolment.

References

1. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001 February 10;322(7282):355-7.
2. Jordhoy MS, Fayers PM, Ahlner-Elmqvist M, Kaasa S. Lack of concealment may lead to selection bias in cluster randomized trials of palliative care. *Palliat Med* 2002 January;16(1):43-9.
3. Borm GF, Melis RJF, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. *Stat Med* 2005 December 15;24(23):3535-47.
4. Philp I. EASYcare: a systematic approach to the assessment of older people. *Geriatric Medicine* 2000;30(5):15-9.
5. Klar N, Donner A. Current and future challenges in the design and analysis of cluster randomization trials. *Stat Med* 2001 December 30;20(24):3729-40.
6. Haidich AB, Ioannidis JP. Patterns of patient enrollment in randomized controlled trials. *J Clin Epidemiol* 2001 September;54(9):877-83.
7. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990 September;65(9 Suppl):S63-S67.
8. Donner A, Birkett N, Buck C. Randomization by cluster. Sample size requirements and analysis. *Am J Epidemiol* 1981 December;114(6):906-14.
9. Slymen DJ, Elder JP, Litrownik AJ, Ayala GX, Campbell NR. Some methodologic issues in analyzing data from a randomized adolescent tobacco and alcohol use prevention trial. *J Clin Epidemiol* 2003 April;56(4):332-40.
10. Moerbeek M, van Breukelen GJ, Berger MP. A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies. *J Clin Epidemiol* 2003 April;56(4):341-50.
11. 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 December;43(11):1601-10.
12. Slymen DJ, Hovell MF. Cluster versus individual randomization in adolescent tobacco and alcohol studies: illustrations for design decisions. *Int J Epidemiol* 1997 August;26(4):765-71.
13. Atherton JS. Teaching and learning: learning curves [on-line]. http://www.learningandteaching.info/learning/learning_curve.htm . 2004.
14. Reynolds B. *Learning and teaching in the practice of social work*. 2nd ed. New York: Russell and Russell; 1965.
15. Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. *JAMA* 2001 November 28;286(20):2578-85.

Chapter 5

Pseudo cluster randomisation performed well when used in practice

R.J.F. Melis, S. Teerenstra, M.G.M. Olde Rikkert, G.F. Borm

J Clin Epidemiol, in press (revised)

Objective: Pseudo cluster randomisation (PCR) randomised researchers in two groups (H and L). Accordingly, patients of one researcher were randomised *in majority to intervention or to control arm. We verified the assumptions basic to the use of PCR in a services evaluation trial: 1) researchers had treatment arm preferences 2) that would have affected recruitment when cluster randomisation had been applied, 3) individual randomisation would have caused contamination, and that PCR prevented 4) allocation predictability, 5) selection bias, and 6) enrolment rate differences.*
Study Design and Setting: Researchers indicated treatment arm preferences, recruitment behaviour in a cluster randomised design, possible contaminating behaviour, and their estimation of the allocation ratio used to allocate their patients. We compared patients' baseline characteristics and individual researchers' recruitment rates.
Results: Researchers favoured the intervention arm (Visual Analogue Scale 14.5 (SD 15.6); 0-100). 58% expected to recruit fewer patients, had every participant been allocated to the control group. 67% estimated that a 50:50 randomisation ratio was used. 65% of researchers used elements of the intervention in control patients. H- and L-researchers recruited a median number of three patients who were comparable at baseline.
Conclusion: The assumptions underlying PCR largely applied in this study.

Introduction

Conventional research methodology does not always satisfy when designing an individual intervention study. While designing a randomised trial on the effects of a problem-based intervention model for community-dwelling elderly people, the authors ran into the limitations of the accepted methods for randomisation [1]. Randomisation at the level of the individual patient was expected to lead to contamination bias, while the accepted solution for contamination bias, cluster randomisation, would probably introduce selection bias and recruitment problems [2, 3]. In order to deal with this dilemma we designed an alternative randomisation method called pseudo cluster randomisation [4, 5]. The aim of this paper is to evaluate the performance of this randomisation procedure.

The principal objective of the services evaluation study in which pseudo cluster randomisation was used was to determine the effects of a nurse-led home visiting programme compared to usual care in improving health-related quality of life in independently living older people with common geriatric problems. In this care programme the general practitioner recruited patients on a problem base.

In this paper the recruiting general practitioner will be further addressed as "researcher", while the physician with recruited patients will be referred to as "cluster". In our study blinding of patients was impossible due to the character of the intervention. The exposure to the intervention resulting from the researcher's participation in the model probably would lead to contamination of control patients, when patients were randomised individually. If we had used cluster randomisation [6-8], the researcher would have known the randomisation decision after the inclusion of the first patient. We anticipated this would have influenced the rate of recruitment and the selection of the patients in the control group [9]. Different selection of patients in the therapy arms would have led to selection bias and incomparable treatment arms.

In the pseudo cluster randomisation method we used, the researchers were first randomised into two types of clusters; clusters H (High) and clusters L (Low). In the

second step randomisation at the patient level was carried out within these clusters. In H-clusters the majority of the subjects received the intervention, while the smaller rest received usual care. In L-clusters the randomisation ratio was reversed.

This approach had important advantages, directed towards selection bias and contamination. The researchers did not know in which cluster group they were, nor did they know in advance what treatment a patient would be on. This reduced the chance of selection bias. No longer half of the researchers were caught in a cluster that was randomised to usual care for all patients. Researchers may be more willing to recruit, when every patient has the prospect of participation in the innovative treatment with expected benefit over regular care.

In L-clusters most of the patients received usual care and only few patients were on the intervention programme. Contamination was expected to be smaller compared to individual randomisation, because there were limited possibilities for the researchers to change their behaviour as a result of contact with the intervention treatment. The controls in H-clusters might be contaminated, as the majority of the patients was on the intervention programme, but the number of controls in these clusters are relatively small.

While these assumptions sound reasonably, they need empirical evaluation. In this paper we will carefully scrutinise the performance of pseudo cluster randomisation in our trial. We will investigate the sustainability of the following underlying assumptions:

- the researchers preferred a certain treatment arm (in which case cluster randomisation would have led to recruitment bias);
- advance knowledge of the outcome of the randomisation would have influenced the researchers' inclusion behaviour (in which case cluster randomisation would have led to recruitment bias);
- the researchers used elements from the intervention treatment in their treatment of control patients or learnt from their involvement in the intervention treatment (in which case individual patient randomisation would have led to contamination);
- the researchers remained blind for the randomisation proportions they were assigned to (i.e. no allocation disclosure in case of pseudo cluster randomisation);
- the two types of clusters included equal numbers of patients (i.e. no recruitment problems in case of pseudo cluster randomisation); and
- the two types of clusters included the same type of patients (i.e. no selection bias in case of pseudo cluster randomisation).

Methods

The evaluation of pseudo cluster randomisation was executed within the setting of a services evaluation study. All recruiting researchers were randomly assigned to one of two groups: clusters H and clusters L. In clusters H recruited patients were randomly assigned using randomisation proportions assigning 80% to the intervention arm and 20% to the control arm. In clusters L recruited patients were randomly assigned using randomisation proportions assigning 20% to the intervention arm and 80% to the control arm.

For the randomisation of the researchers, a minimisation algorithm was used to balance for the factors: 1) high versus low percentage of elderly per practice and 2) the availability of a nurse practitioner.

In the evaluation of the performance of the pseudo cluster method we used of the following assessments:

Baseline characteristics of researchers

We assessed the baseline characteristics of the researchers using a questionnaire researchers filled out before the inclusion of the first patient from their practice. This questionnaire provided data on the following measures of researcher and practice: age, sex, experience, sideline activities, working hours, type and location of practice, number of patients registered at the practice, the proportion of registered older people, and the availability of a practice nurse. The questionnaire also assessed the self-reported competence in geriatrics using the questionnaire of Robinson et al [10].

End of trial characteristics of researchers

When the services evaluation study was ended all researchers received a questionnaire that assessed aspects of their appreciation of the intervention under study, treatment arm preferences when including a participant, their estimation of the randomisation ratio used in their cluster, and the certainty of their estimation, using Visual Analogue Scales. The questionnaire also addressed their alleged inclusion behaviour had they known in advance that all their patients would be in the intervention or in the regular care group, respectively. Finally, the questionnaire investigated once more the self-reported competence in geriatrics using the Robinson questionnaire [10].

Baseline characteristics of patients

We used a written questionnaire and an interview to collect baseline patient data. These instruments provided data on: age, gender, and functional abilities (Groningen Activity Restriction Scale (GARS))[11]. Also, data were collected on cognitive status (Mini Mental State Examination)[12], mobility (Timed Up Go Test)[13] and (health-related) quality of life (MOS-20 [14] , Cantril's self anchoring ladder [15], Dementia Quality of Life [16]). Moreover, we collected data on Socio Economic Status, living conditions, loneliness (Loneliness scale de Jong Gierveld) [17], availability of informal caregiver, and informal caregiver burden (Zarit Burden Interview) [18].

Results

A number of 55 researchers agreed to recruit patients for the services evaluation study. The researchers were 47 years of age on average and had been working a mean of 16 years as a general practitioner. Forty-six% of researchers were male. Many researchers had sideline activities, among which educational activities (69%, mainly supervising general practitioner registrars) and management tasks (15%) were most mentioned. Reasons to participate were multiple; important motivators were the expectation that the intervention model would serve as a supplement to regular primary care (69%) and the expected benefits for their patients (62%). The researchers worked in 36 practices that served a mean number of 4145 patients of whom approximately 10% were 75 years of age and over.

Differences between clusters at baseline

In the group of H-clusters were 28 researchers and 27 researchers were in the group of L-clusters. Researchers in L-clusters were less experienced (13 versus 19 years experience, p = 0.03).

When comparing researchers who actually recruited (n = 40) with researchers who did not recruit (n =15), the differences were more prominent. Non-recruiting researchers were younger (43 versus 48 years of age, p = 0.01), less experienced (11 versus 18 years of experience as general practitioner, p = 0.01) and working more often in practices that employed less general practitioners (p = 0.01).

Recruitment

All researchers together recruited 151 patients of whom 85 (56%) were allocated to the intervention and 66 to regular care. In the H-clusters (n = 28) 20 researchers recruited 85 patients (56%) and in the L-clusters (n = 27) 20 researchers recruited 66 patients (table 1). On average, the H-clusters recruited 4.2 patients and the L-clusters recruited 3.4 patients. The number of patients recruited per cluster varied widely: 10 researchers recruited only one patient, eight researchers recruited two patients, and one researcher recruited 15 patients. Both type of researchers recruited a median number of three patients (p = 0.19). On average, a proportion of 0.77 was allocated to the intervention arm in clusters H, in clusters L this proportion was 0.27.

Differences between patients at baseline

Loneliness and MOS-20 role functioning scores showed statistically significant differences at baseline for patients recruited by H- and L-clusters (table 2). L-cluster patients had better role functioning (MOS-20 subscale) scores with at the same time worse loneliness scores.

table 1 *Recruitment characteristics for all clusters and per cluster type*

	All clusters (n = 55)	H-clusters* (n = 28)	L-clusters* (n = 27)	p
No. of researchers that actually recruited (%)	40 (73)	20 (71)	20 (74)	0.91
No. of included patients (% of all included patients)	151	85 (56)	66 (44)	0.12
No of patients per researcher (mean + sd) (median + IQR)	3.8 (3.2) 3 (1.5 – 5)	4.2 (4.0) 3 (2 – 4)	3.4 (2.3) 3 (1 – 5.5)	0.17

* Clusters had a high (H-clusters) or low (L-clusters) proportion of patients randomised to intervention group

Treatment arm preferences

Researchers strongly favoured inclusion in the intervention arm. The score was 14.5 on a VAS where zero indicates strongly favouring the intervention arm and 100 indicates strongly favouring the regular care arm (table 3). ($p < 0.001$ [one sample t-test comparing the value with a value 50 implicating no effect])

table 2 *Baseline characteristics of patients*

	L-clusters† (n = 66)	H-clusters† (n = 85)	p
Age – yr	82.5 (5.9)	81.9 (6.4)	0.38
Female sex – no. (%)	48 (71.6)	65 (77.4)	0.96
Marital status – no. (%)			0.64
Married	20 (30.3)	22 (26.5)	
Divorced	2 (3.0)	5 (6.0)	
Widow(er), partner deceased	36 (54.6)	47 (56.6)	
Single	8 (12.1)	8 (9.6)	
Living together unmarried	0	1 (1.2)	
Living in home for the aged – no. (%)	9 (13.6)	14 (16.7)	0.99
Informal caregiver – no. (%)	55 (82.1)	59 (70.2)	0.78
Receiving home care – no. (%)	35 (52.2)	42 (50.6)	1.00
Groningen Activity Restriction Scale§	34.2 (8.7)	35.3 (8.1)	0.83
MOS-20 subscale mental health||	54.1 (19.5)	53.0 (19.5)	0.76
MOS-20 subscale physical functioning||	16.7 [0 – 50.0]‡	16.7 [0 – 33.3]‡	0.09
MOS-20 subscale role functioning||	0 [0 – 50]‡	0 [0 – 0]‡	0.01
Cumulative Illness Rating Scale-Geriatrics¶	10.0 (4.5)	10.1 (3.5)	0.89
Timed Up and Go – seconds	15.0 [12.0 – 22.0]‡	17.0 [13.0 – 26.0]‡	0.32
Mini Mental State Examination**	22.4 (5.6)	22.5 (5.9)	0.69
Cantril's self anchoring ladder††	6.0 (1.9)	5.6 (2.3)	0.36
Dementia Quality of Life positive affect‡‡	3.1 (0.7)	3.2 (0.7)	0.17
Dementia Quality of Life negative affect§§	2.7 (0.73)	2.7 (0.6)	0.72
Loneliness Scale De Jong-Gierveld||||	5.7 (3.8)	4.4 (3.1)	0.04

* Values are means (SD) unless otherwise stated

† Clusters had a high (H-clusters) or low (L-cluster) proportion of patients randomised to intervention group

‡ For skewed variables median and interquartile range are presented.

§ Groningen Activity Restriction Scale ranges from 18 to 54, with 18 indicating best score

|| MOS-20 subscales ranges from 0 to 100, with 100 indicating best score

¶ Cumulative Illness Rating Scale-Geriatrics ranges from 0 to 20, with 0 indicating no co-morbidity

** Mini Mental State Examination ranges from 0 to 30, with 30 indicating best score

†† Cantril's self anchoring ladder ranges from 0 to 10, with 10 indicating best score

‡‡ Dementia Quality of Life positive affect ranges from 0 to 5, with 5 indicating best score

§§ Dementia Quality of Life negative affect ranges from 0 to 5, with 0 indicating best score

|||| Loneliness Scale De Jong-Gierveld ranges from 0 to 11, with 0 indicating best score

Alleged inclusion behaviour

We asked the researchers how it would have influenced their recruitment behaviour when they would have known about treatment assignment in advance: all their patients allocated to either intervention or control group. Of the researchers 58% responded that they probably would have included fewer patients than they did when they had been certain every participant was allocated to the control group. Of them 2% indicated they would have recruited more patients (table 3).

The others (40%) would not have changed their inclusion behaviour. Of the researchers 48% researchers indicated they would have recruited more patients when they had known with certainty that all patients were allocated to the intervention. The others (52%) indicated this would not have influenced their recruitment.

table 3 *Treatment arm preferences and alleged inclusion behaviour when researchers would have had prior knowledge on treatment assignment*

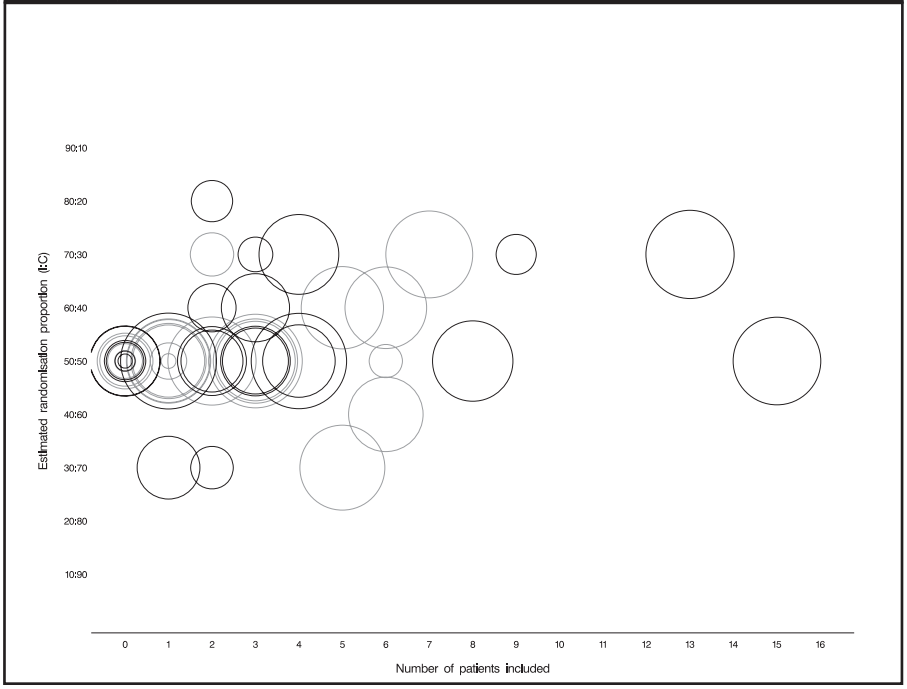
	All researchers (n =55)
Preference for treatment arm when including a participant* – mean (SD)	14.5 (15.6)
Inclusion behaviour when all patients in regular care group – number (%)	
• More	1 (2)
• Less	29 (58)
• No effect	20 (40)
Inclusion behaviour when all patients in intervention group – number (%)	
• More	24 (48)
• Less	0
• No effect	26 (52)

* Visual analogue Scale ranges from 0 = strongly in favour of intervention group to 100 = strongly in favour of control group

Predictability of the next treatment allocation

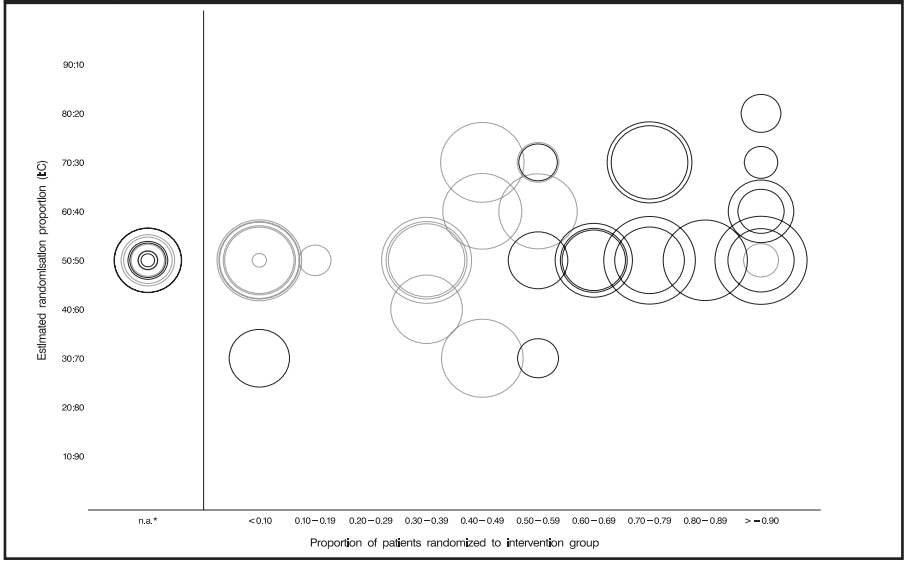
We asked the researchers to estimate which randomisation ratio we used to allocate their patients over intervention and control arm. They could choose nine possibilities ranging from 90:10 / intervention:control arm to 10:90 / intervention:control. There were no differences between the two cluster types, neither with respect to the estimates of the treatment ratios (p = 0.91) nor the degree of certainty (p = 0.65). Of the H-cluster researchers 63% and 71% of the L-cluster researchers indicated that the ratio 50:50

figure 1 *Bubble plot of the relation between the number of inclusions by the researchers and their estimation of the randomisation ratio used to allocate the researcher's patients over the treatment groups*



The grey circles represent researchers randomised to a low number of intervention patients (L-clusters), the black circles represent researchers randomised to a high number intervention patients (H-clusters). The size of the bubble represents the certainty of the estimate; the bubble size decreases with decreasing certainty

figure 2 *Bubble plot of the relation between a researcher's observed randomisation ratio and their estimation of the randomisation ratio used to allocate their patients over the treatment groups*



The grey circles represent researchers randomised to a low number of intervention patients (L-clusters), the black circles represent researchers randomised to a high number intervention patients (H-clusters). The size of the bubble represents the certainty of the estimate; the bubble size decreases with decreasing certainty

*n.a. No randomisation proportion available: of researchers without inclusions no randomisation proportions could be calculated, these researchers were presented separately

was used. Randomisation ratios favouring the intervention arm were indicated in approximately 30% and 20% of the H and L clusters, respectively. About 10% of the researchers in both cluster types indicated randomisation ratios in favour of the control group. The certainty of their estimation was 42.0 on a VAS where zero indicated a complete guess and 100 indicated absolute certainty. Figure 1 shows there was no clear relation between the recruitment numbers and the estimated randomisation ratio. Researchers who estimated randomisation ratios that strongly differ from 50:50, tended to be less certain about their estimation than the researchers who estimated more equal

table 4 *The presence of sources of possible contamination among researchers (n = 23) having patients in both intervention and control arm*

	All clusters (n =55)
Self-perceived effect of having intervention patients on researcher's behaviour when treating control patients – number (%)*	(n = 23)
• Used elements	6 (26)
• Made more referrals to other health care services	7 (30)
• I learnt	9 (39)
• No effect	5 (22)
• Other	1 (4)
Score change in Robinson's self-perceived competence in geriatrics scale – mean (SD)†	6.5 (10.8)

* Only in researchers (n = 23) who had patients in both treatment arms (intervention and control arm)

† Scale ranges from 24 to 120, where the higher score indicates more competence

numbers. Although we found a significant relation between the observed randomisation ratio and the estimated randomisation ratio ($p = 0.02$), the strength of this relation was small: the observed randomisation ratio only explained 15% of the variance in the estimated randomisation ratios ($R^2 = 0.15$) (figure 1).

Indications of potential contamination

Of all researchers having patients in both the control and the intervention arm (n = 23), 15 (65%) indicated that they either used elements of the new intervention, made more referrals to other health care services, or learnt from their participation in the model (table 4). After the study the researchers scored 6.5 points higher on the Robinson competence in geriatrics questionnaire, a scale that ranges from 24 to 120, where a higher score indicates more competence. The score of the recruiting researchers increased almost 8 points, the non-recruiting researchers improved 2.6 points ($p = 0.23$).

Discussion

This study showed that the assumptions underlying the use of pseudo cluster randomisation largely applied when we used this randomisation method in the services evaluation study. We will interpret the results of the current study for each of these assumptions below.

Performance of pseudo cluster randomisation

Assumptions 1 and 2: the researchers preferred to be in the intervention arm of the study and advance knowledge of the randomisation outcome would have influenced the researchers' inclusion behaviour.

Our data show that the researchers in our study strongly favoured allocation of their patients to the intervention arm. This appreciation probably served as an important determinant of the researchers' inclusion behaviour. The majority said they would have included fewer patients when all their patients had been allocated to the control group, as is the case for half of the clusters in a cluster randomised trial. This is in line with other reports that showed that cluster randomisation suffers from lower recruitment rates in the control clusters [9].

Assumption 3: individual patient randomisation would have led to contamination.

Many researchers reported that they used elements of the new intervention in the control group or made more referrals to other health care services for control patients. Also the researchers' self-perceived competence in geriatrics increased. This might point at a learning effect, because this effect that was not present in researchers who did not recruit. These results showed that sources of potential contamination were present in our study and that contamination may have occurred. Individual randomisation would probably have led to even stronger contamination.

Assumption 4: the researchers remained blind for the randomisation proportions they were assigned to.

Most researchers estimated that equal numbers were randomised to both therapy arms. They were not very certain about their estimation and the researchers who estimated an unequal randomisation ratio were even less sure. There was a significant relation between the estimated and observed randomisation proportions. However, the strength of this relation was small, only 15% of the variation in the estimated randomisation

proportions was explained by the observed randomisation numbers. In our study researchers had no or very little knowledge of the randomisation proportions.

Assumption 5 and 6: The two types of clusters included the same number and type of patients. The H-clusters recruited 85 patients and the L-clusters recruited 66. Coincidentally, also 85 and 66 patients were randomised to intervention and control group respectively. There were large differences in the number of patients researchers included. The H-clusters recruited nearly one participant more than the L-clusters. However, the observations were positively skewed and the numbers varied widely. Most researchers included no more than three patients; only two researchers recruited more than 10. Our control group was smaller than the intervention group, but it is unlikely that lack of allocation concealment has caused the difference; researchers were in majority not aware of the allocated randomisation proportions. Patients were comparable at baseline as well, giving no indication of selection bias.

Study limitations

We used the EASYcare trial to evaluate the performance of pseudo cluster randomisation. The best way to evaluate pseudo cluster randomisation would have been to conduct the trial three times: once using individual randomisation, once using cluster randomisation, and once using pseudo cluster randomisation. This was clearly impossible, and we had to settle for weaker sources of evidence such as information on what researchers believed they would have done, had the trial been of a different design. Pseudo cluster randomisation is an innovative design and more studies are required to judge its performance. Nevertheless, pseudo cluster randomisation performed satisfactorily in the services evaluation study.

The assumptions underlying PCR largely applied to the services evaluation study. PCR is a good alternative when individual randomisation risks contamination and cluster randomisation differential recruitment. As such, pseudo cluster randomisation can be added to the range of trial designs investigators can choose from.

References

1. Melis RJF, van Eijken MIJ, Borm GF et al. The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]. *BMC Health Serv Res* 2005 October 5;5(1):65.
2. Jordhoy MS, Fayers PM, Ahlner-Elmqvist M, Kaasa S. Lack of concealment may lead to selection bias in cluster randomized trials of palliative care. *Palliat Med* 2002 January;16(1):43-9.
3. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001 February 10;322(7282):355-7.
4. Borm GF, Melis RJF, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. *Stat Med* 2005 December 15;24(23):3535-47.
5. Teerenstra S, Melis RJF, Peer PG, Borm GF. Pseudo cluster randomization dealt with selection bias and contamination in clinical trials. *J Clin Epidemiol* 2006 April;59(4):381-6.
6. Campbell MK, Grimshaw JM. Cluster randomised trials: time for improvement. The implications of adopting a cluster design are still largely being ignored. *BMJ* 1998 October 31;317(7167):1171-2.
7. Donner A, Birkett N, Buck C. Randomization by cluster. Sample size requirements and analysis. *Am J Epidemiol* 1981 December;114(6):906-14.
8. Donner A, klar N. *Design and analysis of cluster randomization trials in health research*. London: Arnold; 2000.
9. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. *BMJ* 2003 October 4;327(7418):785-9.
10. Robinson BE, Barry PP, Renick N, Bergen MR, Stratos GA. Physician confidence and interest in learning more about common geriatric topics: a needs assessment. *J Am Geriatr Soc* 2001 July;49(7):963-7.
11. 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 December;43(11):1601-10.
12. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of

patients for the clinician. *J Psychiatr Res* 1975 November;12(3):189-98.
13. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 February;39(2):142-8.
14. Kempen GI. [Assessment of health status of the elderly. Application of a Dutch version of the MOS scale]. *Tijdschr Gerontol Geriatr* 1992 August;23(4):132-40.
15. Cantril H. *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press; 1965.
16. Brod M, Stewart AL, Sands L, Walton P. Conceptualization and measurement of quality of life in dementia: the dementia quality of life instrument (DQoL). *Gerontologist* 1999 February;39(1):25-35.
17. de Jong Gierveld J, Kamphuis FH. The development of a Rasch-type loneliness scale. *Appl Psychol Measurement* 1985;9:289-99.
18. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist* 1980 December;20(6):649-55.

Chapter 6

A randomised study of a multidisciplinary programme to intervene on geriatric syndromes in vulnerable older people who live at home (Dutch EASYcare Study)

R.J.F. Melis, M.I.J. van Eijken, S. Teerenstra, T. van Achterberg,
S.G. Parker, G.F. Borm, E.H. van de Lisdonk, M. Wensing,
M.G.M. Olde Rikkert

J Gerontol A Biol Sci Med Sci, in press (revised)

Background: The effectiveness of community-based geriatric intervention models for vulnerable older adults is controversial. We evaluated a problem-based multidisciplinary intervention targeting vulnerable older adults at home which promised efficacy through better timing and increased commitment of patients and general practitioners. This study compared the effects of this new model to usual care.
Methods: General practitioners referred older people for problems with cognition, nutrition, behaviour, mood, or mobility. 151 participants (mean age 82.2 years, 74.8% female) were included in a pseudo cluster randomised trial with six months follow-up for the primary outcomes. Eighty five participants received the new intervention, and 66 usual care. In the intervention arm, geriatric nurses visited patients at home for geriatric assessment and management in cooperation with general practitioners and geriatricians. Modified intention-to-treat analyses focused on differences between treatment arms in functional abilities (Groningen Activity Restriction Scale-3) and mental well-being (subscale mental health MOS-20), using a mixed linear model.
Results: After three months, treatment arms showed significant differences in favour of the new intervention. Functional abilities improved 2.2 points (95% CI 0.3 to 4.2) and well-being 5.8 points (0.1 to 11.4). After six months the favourable effect increased for well-being (9.1 (2.4 to 15.9)), but the effect on functional abilities was no longer significant (1.6 (-0.7 to 3.9)).
Conclusions: This problem-based geriatric intervention improved functional abilities and mental well-being of vulnerable older people. Problem-based interventions can increase the effectiveness of primary care for this population.

Introduction

The autonomy of vulnerable older people is continuously challenged. Chronic diseases, associated functional decline, and erosion of social support systems reduce well-being, and often lead to institutionalisation and high health care costs. Primary health care professionals will care for a substantial part of this expanding group. However, there are significant time limitations in primary care, and there is much room for improvement in quality of geriatric care [1]. This means that developing and evaluating models that enhance primary care for vulnerable older people is an important priority of geriatric primary care research, policy, and practice [2].

Unfortunately, we know little about the effects of geriatric primary care in vulnerable older adults. Critical appraisal of the available evidence is difficult, because the models that can be gathered under the term "community intervention models" show much heterogeneity as well as considerable overlap [3]. We know that preventive home visits can work if they provide multidimensional, high intensity follow-up with clinical control, but there is much debate about effectiveness in vulnerable older people [4, 5]. While some exclude the frailest participants, because of reduced likelihood of reversibility, other authors stress the importance of including the frailest [4-8]. The evidence also suggests that comprehensive geriatric assessment models can work, but this evidence is strongest for inpatient models [9, 10]. Evidence for equivalent community-based interventions is more controversial: in a meta-analysis non-institutional programmes had no effect on hospital re-admission, physical function, or cognitive function [9]. Moreover, most of the included non-institutional programmes are concerned with a general population of older people. The applicability of these results to vulnerable older people is unclear.

A large recent controlled trial of inpatient and outpatient geriatric evaluation and management found significant improvement in mental health of frail persons with outpatient geriatric evaluation and management following a hospitalisation, but not in functional abilities [11].

Case management approaches – a type of care showing overlap with preventive home visiting programmes and comprehensive geriatric assessment – have also been tested in older people [12]. The outcomes are often measures of health care utilisation instead of health outcomes or quality of life [12]. These case management interventions probably have favourable impacts on hospital and long term care utilisation [12, 13], although the recent evaluation of Evercare in the United Kingdom showed no effect on the hospitalisation rate [14]. The study by Bernabei also found beneficial effects on health-related quality of life [15].

In general, the evidence suggests that targeting suitable patients is a key factor in achieving effectiveness [4, 10]. Population screening is a popular approach to targeting [16], but it is expensive and not easy to implement in daily practice. Another important criterion for success is direct involvement of the primary care provider [4].

Therefore, we studied the effectiveness of community intervention models using a problem-based participant selection process performed by the general practitioner. The Dutch Geriatric Intervention Programme (DGIP) is a multidisciplinary community intervention model, consisting of nurse home visits for frail older patients. General practitioners were asked to initiate the intervention when a geriatric condition arose that required further intervention. This procedure promised efficacy through better timing and targeting of the intervention, more engagement of the patient, and more commitment of the general practitioner. The effect of community intervention models for frail older people with this type of targeting has not been rigorously assessed.

In this paper we describe the effects of the Dutch Geriatric Intervention Programme (DGIP) compared to usual care in improving health-related quality of life and promoting successful aging in independently living frail older patients.

Methods

Design

The study design has been published previously [17]. The study was an observer blind, randomised controlled trial that applied pseudo cluster randomisation to allocate the participants to DGIP or usual care [Clinicaltrials.gov Identifier NCT00105378]. The local ethical committee gave approval for the study.

Study population

Subjects lived in their own home or in a retirement home and were 70 years or older (table 1). Participants had one or more limitations in cognition, (instrumental) activities of daily life, or mental well-being.

Randomisation and sample size calculation

Participants were randomised to DGIP or usual care. The usual care group received

table 1 *Eligibility criteria of Dutch EASYcare Study*

Eligibility criteria
Inclusion criteria
• 70 years of age and over
• The patient lives independently or in a retirement home
• The patient has a health problem that was recently presented to the physician by the patient or informal caregiver
• The request for help is related to the following problem fields: cognitive disorders, behavioural and psychological symptoms of dementia, mood disorders, mobility disorders and falling, or malnutrition
• The patient/informal caregiver and physician have determined a goal to achieve
• Fulfil one or more of these criteria: MMSE (Mini Mental State Examination) equal to or less than 26*, GARS-3 (Groningen Activity Restriction Scale) equal to or greater than 25† or MOS-20/subscale mental health equal to or less than 75‡
Exclusion criteria
• The problem or request for help has an acute nature, urging for action (medical or otherwise) within less than one week
• The problem or request for help is merely a medical diagnostic issue, urging for actions only physicians (general practitioner or specialist) can offer
• MMSE < 20 or proven moderate to severe dementia (Clinical Dementia Rating scale [CDR] > 1) and no informal caregiver (no informal caregiver is defined as: no informal caregiver who meets the patient for at least once a week on average)
• The patient receives other forms of intermediate care or health care from a social worker or community-based geriatrician
• The patient is already on the waiting list for a nursing home because of the problem the patient is presented with in our study
• Life expectancy < 6 months because of terminal illness

* MMSE runs from 0-30, with 30 indicating best score

† GARS-3 runs from 18 to 54, with 18 indicating best score

‡ MOS-20 all subscales run from 0 to 100, with 100 indicating best score

unrestricted care. We used a two-step pseudo cluster randomisation procedure, because both individual and cluster randomisation had major drawbacks [18, 19]. Individual randomisation was discarded because it had a risk of contamination bias: the recruiting physician might learn from or use elements of DGIP. However, cluster randomisation would lead to selection bias and lower recruitment rates in the control clusters, because physicians would know the treatment arm their participants would be assigned to after recruiting the first participant [20].

Pseudo cluster randomisation randomised physicians in two groups; group H (high) and group L (low) [18, 19]. The participants recruited through physicians of group H were then randomised in an 80/20 ratio to respectively DGIP and usual care; in group L this ratio was reversed: 20% DGIP and 80% usual care. The physicians were not informed which

group they were in. In the second step of the pseudo cluster randomisation procedure minimisation was used to equally distribute participants for the factors 'high or low percentage of older patients in primary care clinic', 'availability of practice nurse in primary care clinic', 'sex of participant', and 'geriatric condition for referral'. With minimisation the treatment allocated to the next participant enroled in the trial depends on the characteristics of those participants already enroled. This has the advantage, especially in small trials, that there will be only minor differences between groups in those variables used in the allocation process 21. Sample size calculation took account of the pseudo cluster randomised design. We used an uncontaminated minimal detectable difference (MDD) of 4.5 points in the primary outcome measure (GARS-3), with an expected standard deviation of 8.5 (pilot data). We expected that pseudo cluster randomisation would lessen the contamination, though not to the uncontaminated estimate. This means that the MDD has to be set sharper than the uncontaminated level; we used a MDD 4.0 points. Using the usual formula for individually randomised trials with = 0.05 and 1- = 0.80, and a design factor for pseudo cluster randomised trials of 1.08 (cluster size n= 10, intra cluster correlation ρ= 0.05, randomisation fraction f=0.8) (see table 1, Teerenstra et al.[19]) this MDD of 4.0 could be found comparing two groups of 77 patients.

Intervention

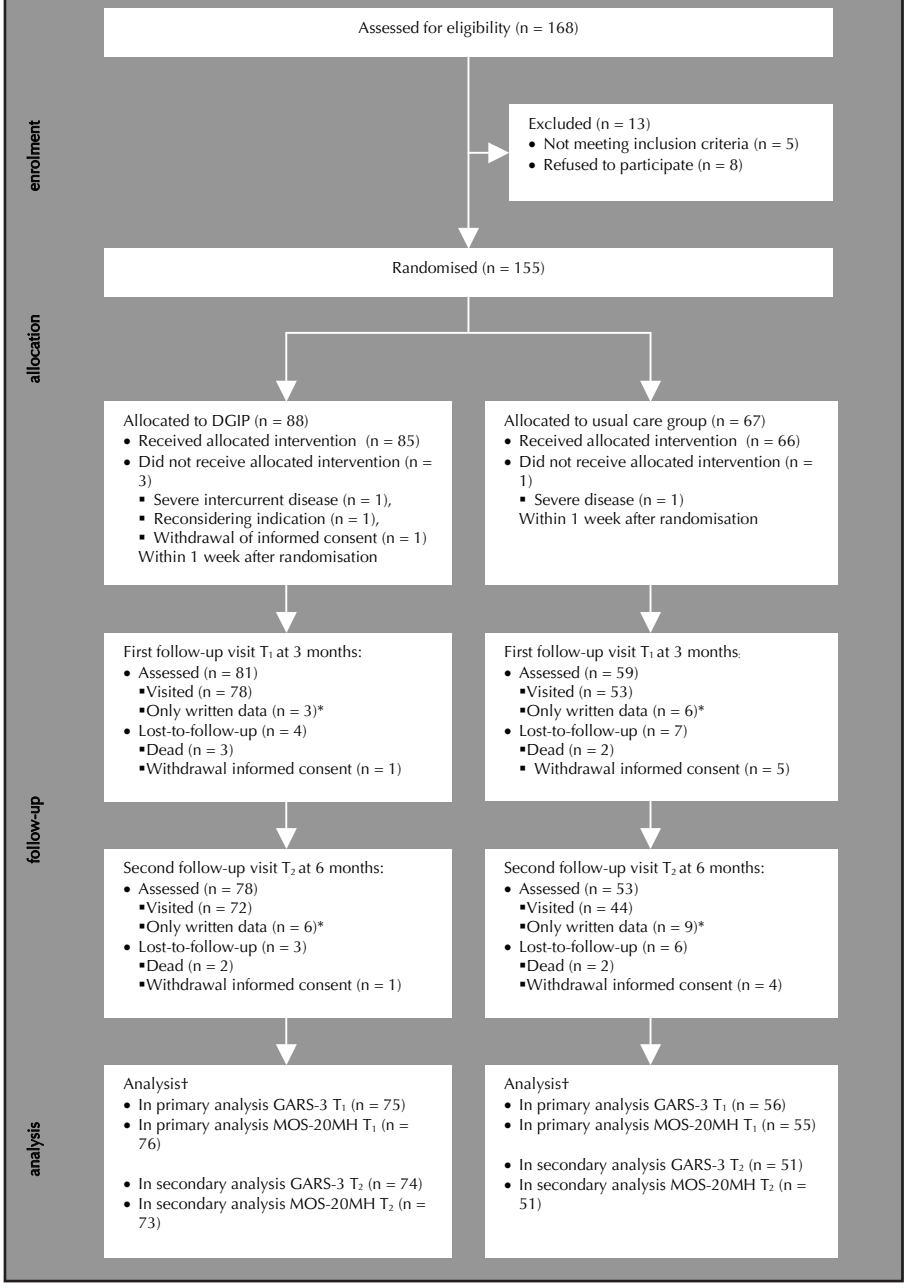
The DGIP used a problem-based selection procedure performed by the general practitioner, rather than population screening to identify patients eligible for participation. The problems targeted concerned cognition, nutrition, behaviour, mood, or mobility, and had to require nursing assessment, coordination of care, therapeutic monitoring, or case management (table 1). Within two weeks after referral a geriatric specialist nurse visited the patient at home. Up to six visits for additional geriatric evaluation and management were planned within the next three months. The nurse conducted the main part of the intervention. The general practitioners continued their usual medical care. Moreover, they made referrals, medication changes, and other interventions as agreed upon during interdisciplinary consultations with the nurse and geriatrician on individual cases. The general practitioner kept prime responsibility for the care of the patient and made the final decisions. We developed guidelines for each of the five presenting health problems to structure activities, without losing the flexibility of tailoring the individual interventions.

Data collection and outcome measures

Researchers (RM, ME), not involved in the conduct of the intervention programme, visited patients at home to obtain written informed consent and to collect baseline (T_0) demographic characteristics and data on general health conditions. If the participant was not able to give informed consent, we asked consent by proxy. Unaware of treatment assignment, the researcher repeated these measurements in the patients' homes three and six months after inclusion. After each follow-up visit the researcher indicated whether blinding remained intact or not.

Primary outcome measures were functional performance in (instrumental) activities of daily living measured using Groningen Activity Restriction Scale-3 (GARS-3) and mental well-being using subscale Mental Health of the Medical Outcome Study 20-item short form (MOS-20 MH) 22. GARS-3 measures 11 basic activities of daily

figure 1 *Study flow chart*



* A number of participants withdrew their consent for the visit for data acquisition by interview, but agreed to fill in the questionnaires.

† Differences between numbers assessed at follow-up and numbers included in the analysis result from the fact that although assessment was executed, participants not always completely filled the instruments and therefore no sum scores could be calculated for everyone who was assessed.

living and 7 instrumental activities of daily living on a three point scale (patient can do activities independently without any difficulty, independently but with difficulty, only with someone's help, respectively). In advance we expected that the larger part of our study population would live in their own home, in whom both types of activities of daily living are very important. Therefore, we used the complete scale as primary measure for functional performance. Secondary outcomes were cognition (Mini Mental State Examination; MMSE 23), mobility (Timed Up and Go test; TUAG 24), loneliness (Loneliness Scale de Jong-Gierveld 25), health-related quality of life (other MOS-20 subscales), Cantril's self anchoring ladder for actual quality of life, Dementia Quality of Life (DQoL) 22, 26, 27, and survival.

Statistical analysis

The primary analysis was a modified intention-to-treat analysis on differences (Intervention – Control) in changes from baseline in GARS-3 and MOS-20MH at three months of follow-up ($T_1 - T_0$). A random effects model was used to account for clustering at the level of the physician 17. The other outcomes at three months follow-up were analyzed in a similar way. Similar analyses at six months follow-up were only performed if the outcome measure showed a significant effect at three months (conditional testing). Kaplan-Meier estimates were used to quantify the intervention effect on survival. We calculated 95% confidence intervals for the differences between treatment arms and used a two-sided alpha of 0.05 to test significance. The baseline characteristics were tested using a random effects logistic model for categorical values and a random effects linear model for continuous outcomes. For skewed variables these models were used with the log transformed scores. Preplanned subgroup analyses of the effects in the primary outcomes – adding the stratifying factor as a covariate and an interaction term of the stratifying factor with treatment arm to the models – were performed for living independently versus living in a retirement home, and higher versus lower levels of cognitive function measured with MMSE (≥ 21 versus <21) at three and six months follow-up.

Results

In and around Nijmegen, the Netherlands, 55 general practitioners participated. During a 21 months inclusion period that started April 1, 2003 155 eligible participants were randomised, of whom four participants did not receive the allocated intervention due to events that took place within one week after randomisation (figure 1). These four participants were excluded from further follow-up and analysis [17]. 85 participants were included in the DGIP-group, and 66 in the usual care group. Mistakenly, one participant aged 69 years was included. This participant was kept in follow-up and analysis. Baseline characteristics and measures of primary outcomes showed no significant differences between study groups. Of secondary outcomes only loneliness differed significantly at baseline (table 2). Our study population mostly comprised widowed females born in the Netherlands, of whom 85% lived on their own. The participants had a mean age of 82.2 years (range 69-99 years), much co-morbidity, MMSE scores suggesting cognitive deterioration, and low scores on mental well-being. Most people had difficulties with all of the (instrumental) activities of daily life measured. Approximately half the study group had home care available at baseline.

table 2 *Demographic characteristics and outcome measures of the study population at enrolment**

	Usual care (n = 66)	Dutch Geriatric Intervention Programme (n = 85)	p
Age – yr	82.8±6.6	81.7±5.9	0.22
Female sex – no. (%)	49 (74.2)	64 (75.3)	0.99
Marital status – no. (%)			0.31
• Married	19 (29.2)	23 (27.4)	
• Divorced	2 (3.1)	5 (6.0)	
• Widow(er), partner deceased	36 (55.4)	47 (56.0)	
• Single	8 (12.3)	8 (9.5)	
• Living together unmarried	0	1 (1.1)	
Living in retirement home – no. (%)	11 (16.9)	12 (14.1)	0.56
Informal caregiver – no. (%)	49 (74.2)	65 (76.5)	0.71
Receiving home care – no. (%)	34 (51.5)	43 (51.1)	0.97
Groningen Activity Restriction Scale-3‡	34.1±8.7	35.3±8.1	0.68
Medical Outcomes Scale-20 mental health§	53.8±17.7	53.3±20.9	0.87
MOS-20 physical functioning§	16.7 [0 – 33.3]†	16.7 [0 – 36.7]†	1.00
MOS-20 role functioning§	0 [0 – 50.0]†	0 [0 – 50.0]†	0.87
Cumulative Illness Rating Scale-Geriatrics||	9.8±4.3	10.2±3.7	0.64
Timed Up and Go – seconds	16.0 [12.0 – 22.0]†	16.0 [12.0 – 25.0]†	0.60
MMSE¶	22.0±6.0	22.8±5.5	0.22
Cantril's self anchoring ladder**	5.9±2.1	5.7±2.1	0.78
Dementia Quality of Life positive affect††	3.1±0.7	3.2±0.7	0.50
DQoL negative affect‡‡	2.6±0.6	2.7±0.7	0.24
Loneliness Scale De Jong-Gierveld§§	5.7±3.6	4.4±3.2	0.04

* For all – except skewed – variables means ± SDs are presented, for skewed variables (†) median and interquartile range are presented.

‡ GARS-3 runs from 18 to 54, with 18 indicating best score

§ MOS-20 all subscales run from 0 to 100, with 100 indicating best score

|| CIRS-G runs from 0 to 20, with 0 indicating no co-morbidity

¶ MMSE runs from 0 to 30, with 30 indicating best score

** Cantril's self anchoring ladder runs from 0 to 10, with 10 indicating best score

†† DQoL positive affect run from 0 to 5, with 5 indicating best score

‡‡ DQoL negative affect run from 0 to 5, with 0 indicating best score

§§ Loneliness Scale De Jong-Gierveld runs from 0 to 11, with 0 indicating best score

The participating general practitioners cared for a mean 1719 patients (SD 470) of whom 170 (SD 131) were 75 years and older. Of this subgroup of older subjects 3% (SD 4) was included in this study. About 40% of participants were referred because of a problem relating to cognition. Mood and mobility problems were reason for referral both in 20% of the cases. Behavioural and nutritional problems were referral reasons in 11 and 6% respectively. The nurse visited intervention patients 3.8 times (SD 1.3). Problem analysis was an important component of these visits. The interventions focused mainly on therapeutic advices and coordination of care, less interventions focused on psychoeducation or therapy monitoring (table 3).

table 3 *Content of the Dutch Geriatric Intervention Programme (DGIP) in intervention group patients (n=85)*

Type of activity	Number (%)
Diagnostic tests	
EASYcare screening	85 (100)
Focused history taking and caregiver assessment	85 (100)
Mini Mental State Examination	40 (47)
Geriatric Depression Scale	30 (35)
Montgomery-Åsberg Depression Rating Scale	14 (16)
Body weight	23 (27)
Mini Nutritional Assessment – Short Form	2 (2)
Body Mass Index	1 (1)
Investigation of calorie- and fluid intake	3 (4)
Tinetti Balance and Mobility Scale	11 (13)
Blood pressure	4 (5)
Blood pressure spinal and standing position	21 (25)
Pulse	14 (16)
Visual acuity using Snellen chart	1 (1)
Mean number of diagnostic tests per case (SD; range)	2.0 (1.8; 0 – 9)
Interventions – care coordination	
(Extra home) care	34 (40)
Consultation of other health care professional	31 (36)
Organise programme for daily structure	21 (25)
Mobilise voluntary care	25 (29)
(Extra) day care	12 (14)
Interventions – advices	
Advices to professional handling behavioural difficulties	18 (21)
Advices to caregiver handling behavioural difficulties	28 (33)
Advice to patient	45 (53)
Referral to other health care professional	24 (28)
Medication change	22 (26)
Start using helping aid (cane etc.)	16 (19)
Interventions – psychoeducation	
Patient	23 (27)
Caregiver	15 (18)
Intervention – monitoring the effect of therapy	26 (31)
Mean number of interventions per case (SD; range)	4.6 (2.3; 0 – 10)

After three months of follow-up the primary outcomes functional performance and mental well-being showed significant treatment arm differences in changes from baseline (table 4). On GARS-3 this difference was -2.2 [95% confidence interval -4.2 – -0.3] and on MOS-20MH it was 5.8 points [0.1–11.4] both in favour of DGIP [table 3]. At six months of follow-up favourable effects still existed, although the effect on GARS-3 was slightly smaller and no longer significant: -1.6 [-3.9 – 0.7]. The effect on MOS-20MH increased to 9.1 [2.4 – 15.9]. In the usual care group the GARS-3 scores worsened from baseline, while during the first three months this decline is absent in DGIP (table 4). DGIP improved MOS-20MH scores over six months follow-up. The MOS-20MH scores remained approximately constant in the usual care group.

table 4 *The unadjusted scores on primary outcome measures at 3 and 6 months and the differences between treatments arms in outcome measures' change from baseline*

	3 months follow-up (T_1)			6 months follow-up (T_2)*		
Primary outcomes unadjusted scores	Usual care	DGIP†		Usual care	DGIP†	
Functional abilities – GARS-3	36.4 ± 10.3	34.7 ± 8.1		37.0 ± 9.5	35.9 ± 8.6	
Well-being – MOS-20MH	55.5 ± 18.2	60.0 ± 20.5		53.2 ± 20.1	61.5 ± 20.4	
Intervention effect‡	**D**	**95% CI**	**N**	**D**	**95% CI**	**N**
GARS-3§	**-2.2||**	**-4.2 – -0.3**	**131**	**-1.6**	**-3.90 – 0.7**	**125**
MOS-20MH§	**5.8||**	**0.1 – 11.4**	**131**	**9.1¶**	**2.4 – 15.6**	**124**
DQoL Negative affect	-0.21||	-0.37 – -0.04	114	-0.13	-0.32 – 0.06	103
DQoL Positive affect	0.15	-0.06 – 0.37	112			
MOS-20 physical performance	4.3	-2.9 – 11.2	132			
MOS-20 role functioning	4.7	-9.8 – 19.3	131			
Cantril's Ladder	0.39	-0.26 – 1.03	105			
Loneliness de Jong-Gierveld	0.10	-0.80 – 0.99	109			
MMSE	-	-	0	-0.5**	-1.8 – 0.1	113
Timed Up And Go	1.60	-4.82 – 8.03	107			

* Analysis per outcome measure at 6 months follow-up were only performed for the condition that the treatment arm difference calculated for 3 months was significant

† DGIP = Dutch Geriatric Intervention Programme

‡ Intervention effect: difference (D) and its 95% Confidence Interval (CI) in changes over time from baseline to three and six months, for the number of subjects (N) in analysis.

Model: Outcome Measure Change from baseline = $\beta_0 + \beta_1$*treatment arm + β_2*baseline value Outcome Measure + β_c*minisation factors + random intercept physician

§ Primary outcomes

|| $P < 0.05$

¶ $P < 0.01$

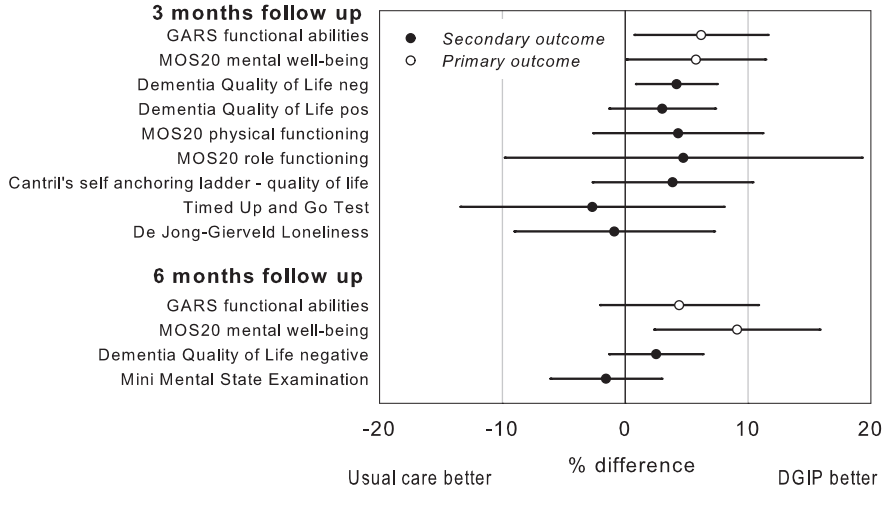
** MMSE was only measured at baseline and at 6 months follow-up

Secondary outcome measures Dementia Quality of life Positive and Negative Affect, subscales Physical Performance and Role Functioning of MOS-20, and Cantril's Ladder showed a trend towards beneficial effects for DGIP. The effects on Mini Mental State Examination, De Jong-Gierveld Loneliness Scale, and Timed Up and Go test were close to zero (figure 2). DGIP survival at two years follow-up was higher (82% versus 73%, Log-Rank test p=0.40).

The results of a sensitivity analysis with loneliness score – the only baseline characteristic that differed between our treatment arms – added as covariate, were in line with the primary analysis.During the follow-up measurements treatment assignment was revealed to the researcher in 38% of cases at three months of follow-up and in 40% of cases at six months of follow-up.

The total drop out in our study was 7% at three months and 13% at six months, and similar in both groups. Participants who were lost-to-follow-up were older and had worse GARS-3, MOS-20MH, and MMSE scores at baseline. The results of sensitivity analyses assigning the 'mean of the other group' [28] to the missing values did not differ essentially from the primary analyses. No significant statistical interactions with MMSE scores or living conditions were found.

figure 2 *Primary and secondary outcomes differences between study arms at three and six months as percentages of scale ranges and their 95% confidence intervals (for Timed Up and Go Test a denominator of 60 seconds was used, because Timed Up and Go Test has no scale range).*



Discussion

This randomised controlled trial found benefits of a nurse-led, multidisciplinary intervention at home on frail older subjects' functional performance and mental well-being at three months of follow-up. At six months the well-being scores had further improved and, the performance on functional abilities, while still better in the intervention group, had not further increased. Most secondary outcomes showed a trend towards advantageous effects. The results of the economic evaluation that accompanies this study showed that this intervention is an effective addition to primary care for frail older people at a reasonable cost 29.

The age, co-morbidity, and GARS-3 scores at baseline and the overall mortality during follow-up, show that a group of very old, vulnerable patients was sampled. The beneficial effects on disability and mental well-being represent a 5% and 10% better performance compared to control conditions, respectively.

The results of our study show that it was possible to prevent deterioration of functional skills for about three months and to improve well-being for at least half a year in a vulnerable population with a fairly simple home-based intervention. The magnitude of these effects is in line with treatment effects of other positive studies incorporating frail populations 11, 15. An evaluation of outpatient geriatric evaluation and management found favourable differences in mental health and physical functioning scores of about 5% and 2% at 12 months of follow-up respectively 11. A trial with a model of integrated care and case management for frail older people living in the community found favourable differences of 18.1% in basic activities of daily living, 6.9% in instrumental activities of daily living, and 6.8% in depression 15. Also, our results were above a (standardized) effect size of 0.2, which is considered to be the lowest threshold for a minimal clinically important difference 30.

Our control group was smaller than the intervention group, but it is unlikely that lack of allocation concealment has caused the difference; physicians were in majority not aware of the allocated randomisation proportions. Patients were comparable at baseline as well, giving no indication of selection bias. An explanation is offered by the variation in the number of patients each physician included. Two physicians included more than 10 patients and both were assigned to the group of general practitioners whose patients were randomised to the intervention group in majority. This observation completely explains the unbalanced numbers of control and intervention group.

The total drop out in our study was fairly high, but was similar in both groups and as expected when taking into account the frailty of the population. Drop outs occurred mainly because patients (or their caregivers) felt participation in follow-up visits for effect measurement was too burdensome while it provided no further benefit. Participants who were lost-to-follow-up differed from participants who completed follow-up. However, the results of sensitivity analyses – using a conservative strategy to impute missing values 28 – did not differ essentially from the primary analyses.

This study was observer blind. Despite several precautionary measures taken, disclosure of treatment assignment occurred frequently. However, our primary outcomes were collected using a written questionnaire that the patient (if necessary with help from a relative) filled before each study visit. The researcher could not influence this.

Given the type of patients included in this study, the study results can probably be generalized to a population of frail community-dwelling older people. General practitioners appeared to be very selective. Approximately 3% of all older patients cared for by one general practitioner were included in this study. However, we have to keep in mind that only a minority of older patients can be characterized as vulnerable, depending on the definition 31. This means that only a minority actually is eligible for this intervention that explicitly focused on frail subjects who also needed to have an incident geriatric problem. Unfortunately, we were unable to collect further details on the patients who were not included. This means that generalisation of these results to the general population of community-based older persons deserves further evaluation. However, generalisation benefits from the fact that study conditions were very similar to current practice. Even without much experience with the model, the general practitioners were able to select patients eligible for the intervention. As under regular conditions, the nurses had to cooperate with many different health care workers. The results of our trial show that multidimensional intervention for geriatric syndromes improves disability and mental well-being in frail older people who live at home. The results also indicate that this can be done by the general practitioner using a problem-based patient selection procedure to target suitable patients. As such, it promises to be a relevant supplement to primary health care for this population. This is important because population aging and increasing awareness of patient autonomy will increase the number of frail older people who rely on primary health care in reaching the aim of successful aging.

References

1. Wenger NS, Solomon DH, Roth CP et al. The quality of medical care provided to vulnerable community-dwelling older patients. *Ann Intern Med* 2003 November 4;139(9):740-7.
2. Reuben DB. Meeting the needs of disabled older persons: can the fragments be pieced together? *J Gerontol A Biol Sci Med Sci* 2006 April;61(4):365-6.
3. Parker G, Bhakta P, Katbamna S et al. Best place of care for older people after acute and during subacute illness: a systematic review. *J Health Serv Res Policy* 2000 July;5(3):176-89.
4. Rubenstein LZ, Stuck AE. Preventive home visits for older people: defining criteria for success. *Age Ageing* 2001 March;30(2):107-9.
5. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. *JAMA* 2002 February 27;287(8):1022-8.
6. Egger M. Commentary: When, where, and why do preventive home visits work? *BMJ* 2001;323(7315):724-5.
7. Elkan R, Kendrick D, Dewey M et al. Effectiveness of home based support for older people: systematic review and meta-analysis. *BMJ* 2001 September 29;323(7315):719-25.
8. van Rossum E, Frederiks CM, Philipsen H, Portengen K, Wiskerke J, Knipschild P. Effects of preventive home visits to elderly people. *BMJ* 1993 July 3;307(6895):27-32.
9. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993 October 23;342(8878):1032-6.
10. Wieland D. The effectiveness and costs of comprehensive geriatric evaluation and management. *Crit Rev Oncol Hematol* 2003 November;48(2):227-37.
11. Cohen HJ, Feussner JR, Weinberger M et al. A controlled trial of inpatient and outpatient geriatric evaluation and management. *N Engl J Med* 2002 March 21;346(12):905-12.
12. Johri M, Beland F, Bergman H. International experiments in integrated care for the elderly: a synthesis of the evidence. *Int J Geriatr Psychiatry* 2003 March;18(3):222-35.
13. Beland F, Bergman H, Lebel P et al. A system of integrated care for older persons with disabilities in Canada: results from a randomized controlled trial. *J Gerontol A Biol Sci Med Sci* 2006 April;61(4):367-73.
14. Gravelle H, Dusheiko M, Sheaff R et al. Impact of case management (Evercare) on frail elderly patients: controlled before and after analysis of quantitative outcome data. *BMJ* 2007 January 6;334(7583):31.
15. Bernabei R, Landi F, Gambassi G et al. Randomised trial of impact of model of integrated care and case management for older people living in the community. *BMJ* 1998 May 2;316(7141):1348-51.
16. Wieland D, Rubenstein LZ. What do we know about patient targeting in geriatric evaluation and management (GEM) programs? *Aging (Milano)* 1996 October;8(5):297-310.
17. Melis RJF, van Eijken MIJ, Borm GF et al. The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]. *BMC Health Serv Res* 2005 October 5;5(1):65.
18. Borm GF, Melis RJF, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. *Stat Med* 2005 December 15;24(23):3535-47.
19. Teerenstra S, Melis RJF, Peer PG, Borm GF. Pseudo cluster randomization dealt with selection bias and contamination in clinical trials. *J Clin Epidemiol* 2006 April;59(4):381-6.
20. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? *BMJ* 2001 February 10;322(7282):355-7.
21. Altman DG, Bland JM. Treatment allocation by minimisation. *BMJ* 2005 April 9;330(7495):843.
22. Kempen GI. The MOS Short-Form General Health Survey: single item vs multiple measures of health-related quality of life: some nuances. *Psychol Rep* 1992 April;70(2):608-10.
23. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975 November;12(3):189-98.
24. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc* 1991 February;39(2):142-8.
25. de Jong Gierveld J, Kamphuis FH. The development of a Rasch-type loneliness scale. *Appl Psychol Measurement* 1985;9:289-99.
26. Brod M, Stewart AL, Sands L, Walton P. Conceptualization and measurement of quality of life in dementia: the dementia quality of life instrument (DQoL). *Gerontologist* 1999 February;39(1):25-35.
27. Cantril H. *The pattern of human concerns*. New Brunswick, NJ: Rutgers University Press; 1965.
28. Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med* 2001 December 30;20(24):3931-46.
29. Melis RJF, Adang E, Teerenstra S et al. The cost-effectiveness of a multidisciplinary intervention model for community-dwelling frail older people. J Gerontol A Biol Sci Med Sci. In press 2007.
30. Cohen J. A Power Primer. *Psychological Bulletin* 1992 July;112(1):155-9.
31. Fried LP, Ferrucci L, Darer J, Williamson JD, Anderson G. Untangling the concepts of disability, frailty, and comorbidity: implications for improved targeting and care. *J Gerontol A Biol Sci Med Sci* 2004 March;59(3):255-63.

Chapter 7

The effect on caregiver burden of a problem-based home visiting programme for frail older people in a randomised study

R.J.F. Melis, M.I.J. van Eijken, T. van Achterberg, M. Vernooij-Dassen, E.H. van de Lisdonk, M.G.M. Olde Rikkert

Submitted

Objectives: Caregiver effects of geriatric care models focusing primarily at the patient have not been consistently studied. We studied caregiver effects of a nurse-led comprehensive geriatric evaluation and management (GEM) programme for community-dwelling frail older people in comparison with usual care.
Methods: This randomised trial included 110 caregiver/patient couples who were followed for six months. Primary analyses were intention-to-treat analyses of caregiver burden assessed with Zarit Burden Interview (ZBI; 0-88; higher means more burden). Preplanned subgroup analyses were conducted for cognition, living conditions, and household composition.
Results: Overall perceived caregiver burden showed no significant differences between study groups in changes over time. However, perceived burden was at baseline more than eight points higher in caregivers sharing a household with patients (n = 23) compared to caregivers living separately (n = 87). The intervention performed convincingly better in caregivers living together with the patient than in caregivers living separately (p for interaction = 0.04).
Discussion: GEM caregiver effects might be related to whether caregivers live with the patient or not. Baseline differences in perceived burden in both subgroups may offer an explanation: it is the most burdened group, sharing households with the patient, which benefits from the intervention.

Introduction

Increasingly, community-based alternatives for institutional geriatric care are disseminated throughout health care [1, 2]. There are very legitimate reasons for this up rise of primary care geriatric evaluation and management (GEM) models: as result of population aging and increasing awareness of patient autonomy more and more frail older people will rely on primary health care 3, 4. However, despite the promises this type of health care holds for vulnerable older people, these models will evidently affect the family and friends who care for them. This influence is not necessarily only positive. Caregivers are in a special position, because on the one hand they are supporting the patient, whereas on the other they are in need of support for themselves 5. Although some of these models have an eye for the caregivers and their burden, successful intervention in patients often relies on the caregivers' participation as well. This may increase the burden of caring.

Many studies have evaluated the effect of interventions primarily aimed at caregivers 6-9. These studies suggest that interventions aimed at that caregiver improve outcomes such as caregiver burden, depression, or subjective well-being. This is true both when the care recipient is an older adult in general as well as when the recipient has a cognitive disorder [8, 9]. The increasing interest in dyadic modelling of the caregiving relationship shows how scientists increasingly appreciate the complexity and reciprocity of the relationships between caregivers and the patients. [10]. This also affects how caregiver interventions are being evaluated. The focus has shifted from solely caregiver outcomes to also paying attention to how care recipients are affected in interventions focused on the caregivers. Despite the increasing attention given to the care recipient or patient in caregiving research, not many have studied this relationship while evaluating models aimed at the older patient in the first place [11]. This may be the consequence of the intuitive expectation that improving the life of the patient also protects the informal caregiver from the increasing burden that often accompanies caring for a frail, aging person. However, these interventions could also increase the caregiver's burden by recruiting even more of the caregivers assistance [11]. These potential 'side-effects' of GEM models are often neglected in GEM evaluations, but have been evaluated in a few studies

[11-14]. These studies suggest positive effects of outpatient GEM on caregiver burden. The Dutch Geriatric Intervention Programme (DGIP) is a comprehensive geriatric assessment model, consisting of nurse home visits for frail older patients, that has shown to effectively improve functional abilities and well-being of frail older patients at a reasonable cost [15]. The objective of this paper is to describe the effects of this model of home visits that focuses primarily at the community-dwelling frail older people themselves on caregiver burden in comparison with usual care. Based on the limited findings available, we hypothesised that the intervention would have a beneficial effect on caregiver burden.

Methods

Design

We conducted an observer blind randomised controlled trial of which the design was previously published [16]. The local Ethical committee approved of the study.

Study population

General practitioners referred independently living older patients to our model when there was a problem in cognition, nutrition, behaviour, mood, or mobility. The problem had to urge for nursing assessment, coordination of care, or therapeutic monitoring and case management. Referrals for acute problems or purely medical issues were not accepted. Patients had to live in their own home or in a home for the aged and should be 70 years of age or older. An informal caregiver was defined as a person who felt responsible for the care of the patient and who was the closest involved in that care as non-professional.

We restricted patient inclusion to those patients with one or more limitations in cognition, (instrumental) activities of daily life, or mental well-being. The exact cut-offs were on Mini Mental State Examination (MMSE): equal to or less than 26, on Groningen Activity Restriction Scale (GARS-3): equal to or greater than 25, and on MOS-20 subscale mental health (MOS-20MH): equal to or less than 75 [17-19]. Moreover we excluded persons already receiving forms of intermediate care or health care from a social worker or community-based geriatrician, persons on the waiting list for a nursing home, and those who had a life expectancy of less than six months, because of terminal illness.

Randomisation

Patients/caregiver couples were randomised over two treatments arms: home visits and usual care. The usual care group received unrestricted care. We used a two-step pseudo cluster randomisation procedure; because individual randomisation risked contamination and cluster randomisation risked selection bias due to expected treatment arm preferences of recruiting physicians. We reported details of this novel design elsewhere [16, 20, 21]. Pseudo cluster randomisation divided the general practitioners in two groups; a group H and a group L. The patients recruited through physicians of group H were randomised in an 80/20 ratio to home visits and usual care; while in group L this ratio was reversed: 20% home visits and 80% usual care. The occurrence of selection bias and poor recruitment is prevented, because treatment allocation remains concealed. Limited exposure of L-physicians to the home visits programme lowers risk of contamination of usual care.

Intervention

The intervention (Dutch Geriatric Intervention Programme; DGIP) used a problem-based selection procedure performed by the general practitioner. The problems targeted concerned cognition, nutrition, behaviour, mood, or mobility, and had to require nursing assessment, coordination of care, therapeutic monitoring, or case management. Within two weeks after referral a geriatric specialist nurse visited the patient at home. Up to six visits for additional geriatric evaluation and management were planned within the next three months. The nurse conducted the main part of the intervention. The general practitioners continued their usual medical care. Moreover, they made referrals, medication changes, and other interventions as agreed upon during interdisciplinary consultations with the nurse and geriatrician on individual cases. The general practitioner kept prime responsibility for the care of the patient and made the final decisions. We developed guidelines for each of the five presenting health problems to structure activities, without losing the flexibility of tailoring the individual interventions. Although not a standard intervention, the nurse also performed a caregiver burden assessment [22, 23]. The results were implemented in the care plan; the nurse gave for instance extra attention to psycho-education, organised extra adult day care, or taught the caregiver how to handle behavioural disturbances in dementia. Caregivers were involved in the conception and execution of the intervention to their capability. The nurse, geriatrician, and general practitioner had frequent consultations on individual cases.

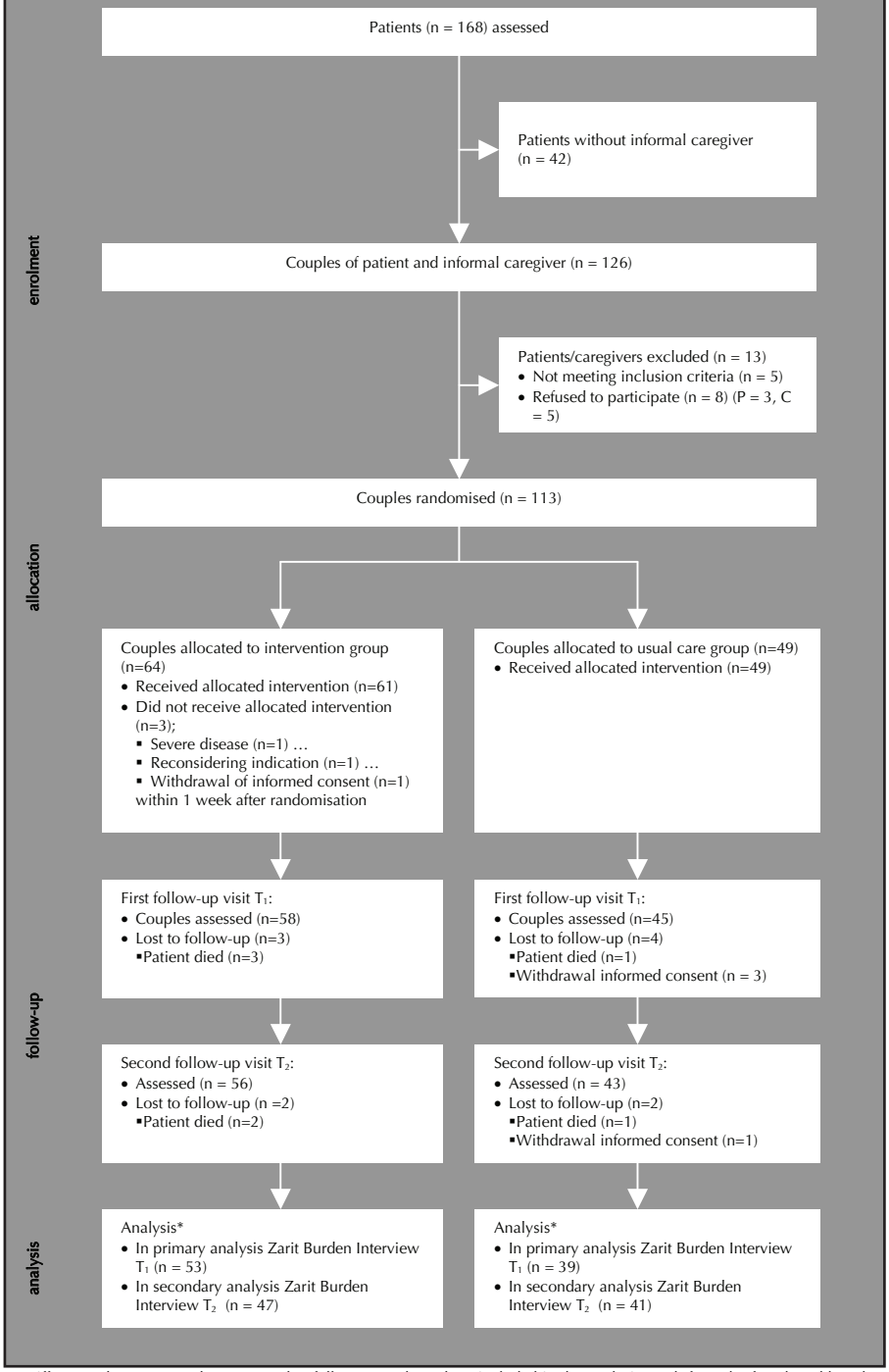
Data collection and outcome measures

A researcher interviewed patients at home to obtain written informed consent and to collect baseline demographic characteristics, data on general health conditions, and data on use of home care. If the patient was not capable to give informed consent, we asked a proxy to do so. We collected data on age and sex of the caregiver, their relationship with the patient, type and amount of care provided, time spent on caring, and caregiver burden. Also data were collected on competing time demanding activities such as work, caring for a family, or leisure time activities.These measurements were repeated three and six months after inclusion. The primary outcome measure for the caregivers was caregiver burden measured with the Zarit Burden Interview (range 0-88, with burden increasing with the score). The secondary outcome is time spent on care.

Statistical analysis

The baseline characteristics were tested using a mixed logistic model for categorical values and a mixed linear model for continuous outcomes. The primary analysis was an intention-to-treat analysis on the difference in the change from baseline ZBI to three months of follow up, which is at the end of the nursing intervention. A mixed linear model was used to account for clustering at the level of the general practitioner with the baseline measurement of ZBI and the stratification factors used in the randomisation as covariates. In a secondary analysis the difference in change from baseline at six months using the same model was evaluated. The other outcome measures were analysed in a similar way. We calculated 95% confidence intervals for the differences between treatment arms and used a two-sided alpha of 0.05 to test statistical significance. Subgroup analyses will be performed for living independently versus living in a home for the aged, higher versus lower levels of cognitive function, and patient and caregiver living together or not.

figure 1 *Study flow chart of patient/caregiver couples participating in our study*



*Differences between numbers assessed at follow-up and numbers included in the analysis result from the fact that although assessment was executed, participants not always completely filled the instruments and therefore no sum scores could be calculated for everyone who was assessed.

Results

In and around Nijmegen, the Netherlands, 55 general practitioners participated. During an inclusion period of 21 months that started at April 1, 2003, 155 patients eligible for the study were included as patients (Figure 1). In 113 of 155 patients an informal caregiver was identified. Three patient/caregiver pairs were not included in follow up and analysis because these patients experienced severe intercurrent disease necessitating hospital admittance, died, or withdrew informed consent within one week after randomisation. Therefore 110 patient/caregiver couples were included in follow up and analysis; 61 were included in the intervention group and received home visits, 49 were included in the regular group. The caregivers who participated in this trial had a mean age of 56.3 years and 73% of them were females (Table 1). They cared for very frail patients, who were mostly females (72%) with a mean age of 82.2 years and who had much co-morbidity (table 2). Fifty-eight percent of the caregivers were daughters or sons of the patient. Nineteen percent of caregivers were spouses and 21% of caregivers lived together with the patient. Caregivers who did not share their household with the patient lived median only 5 kilometers away, and their travel time was approximately 10 minutes by car. Patients experienced episodes of anxiety, wandering, or confused, restless or panicking behaviours, according to 96% of the caregivers. More than half the caregivers (52%) felt they always had to be watchful or could not leave their relative safely. Caregivers and patients spent a median six hours per week together. During this time the caregivers spent half an hour on personal care and two hours on assistive activities such as administration, organizing care, domestics etc. Another two hours were spent just visiting. These data were skewed; some caregivers spent more than a working week caring for their relative.

Of the caregivers 11% indicated they could rely on no one else in the care for the patient. Offspring caregivers often shared caring for their parent with their siblings (44 of 63 offspring caregivers) or their own spouses (9 of 63), spouse caregivers shared the caring with their children (12 of 20 spouse caregivers) or had no one else to help (6 of 20). Sixty-two percent of the caregivers spent less time on other activities because of the care for the patient. Care for someone else than the patient (13 %), and leisure time activities (23 %) were most mentioned as activities they spent less time on. At baseline caregivers scored a mean ZBI score of 29.9. The age of the patient was higher in the usual care group (Table 1).

About 40% of patients were referred because of a problem relating to cognition. Mood and mobility problems were reason for referral both in 20% of the cases. Nutritional and behavioural problems were referral reasons in 6 and 11% respectively. The nurse visited intervention patients 3.8 times (SD 1.3). These visits lasted 1.3 hours on average. Problem analysis was an important component of these visits. The interventions focused mainly on advices and coordination of care. In 28 cases the nurses taught the caregiver how to handle a specific behaviour and in 15 cases they educated the caregiver.

After three months of follow up the burden of caregivers in the intervention groups showed a non-significant improvement of 0.67 points [95% confidence interval -4.03 – 2.69] compared to the change in burden over this period in the control caregivers. After six months the burden of caregivers in the intervention group showed a non-significant increase of 2.29 points [-1.61 – 6.19] compared with the change over this time period

table 1 *Demographic characteristics and outcome measures of the study population at enrolment*

	Usual care (n = 49)	Home visits (n = 61)	P-value
PATIENT			
Age patient, mean ± SD	83.5 ± 6.1	81.1 ± 5.7	0.03
Female sex patient, n (%)	35 (71.4)	44 (72.1)	0.98
Patient living in a home for the aged, n (%)	9 (18.8)	6 (9.8)	0.24
Mini Mental State Examination patient*	21.1 ± 5.6	21.9 ± 5.8	0.51
CAREGIVER			
Age caregiver, mean ± SD	56.3 ± 13.0	55.5 ± 12.6	0.52
Female sex caregiver, n (%)	33 (68.8)	46 (75.4)	0.39
Living together with the patient, n (%)	7 (14.3)	16 (26.2)	0.17
Travel distance (km), if living separately, median [Interquartile range]	4.7 [2.1 – 13.0]†	5.1 [2.4 – 9.3]†	0.21‡ §
Travel time by cars (min), if living separately, median [Interquartile range]	10.0 [5.0 – 21.0]†	11.0 [6.5 – 18.0]†	0.15‡ §
Relationship with patient, n (%)			0.08
• Spouse	5 (10.2)	15 (25.4)	
• Son/daughter	30 (61.2)	33 (55.9)	
• Other	14 (28.6)	11 (18.6)	
Time spent… (hours/week), median [Interquartile range]			
• Together	5.5 [3.5 – 10.0]†	6.0 [3.5– 168.0]†	0.06‡
• On personal care	0.4 [0 – 2.0]†	0.5 [0 – 3.0]†	0.83‡
• On assistive activities: housekeeping, administration, etc.	2.0 [1.0 –4.0]†	3.0 [1.0 – 7.0]†	0.15‡
Had no help from others in caring, n (%)	6 (12.2)	9 (14.8)	0.76
Had less time for other activities because of care for patient, n (%)	26 (53.1)	33 (54.1)	0.96
Zarit Burden Interview, mean ± SD§	30.4 ± 14.5	27.9 ± 12.4	0.33

* MMSE ranges from 0 to 30, with 30 indicating best score.

† For skewed variables median and quartiles are presented.

‡ For skewed variables P-values were calculated using the square root transformed score.

§ Zarit burden interview ranges from 0 to 88, with zero indicating best score.

in control caregivers (Table 3).

The intervention increased the time spent on other assistive activities such as administration and housekeeping with 1.9 hours/week [-2.2 – 6.0] at six months follow up, but this increase was non-significant. No effect on the amount of time spent on personal care was found.

The burden of caregivers in the usual care group was approximately constant over time, in the intervention group burden tended towards an increase (Table 4). Caregivers sharing a household with their relative (n = 23) were older (mean age 73 vs 52 years for caregivers living separately), more frequently men (11 of 23 vs 16 of 75), almost always spouse (20 of 23 vs 0 of 85), and had higher ZBI scores at baseline (mean score 33,9 (SD 15.0) vs 27.9 (12.8)) than caregivers living separately. Also caregivers sharing a household spent more time on personal care (median 7.0 vs 0.25 hours/week) and more time on other assistive activities (median 10.5 vs 2.0 hours/week) at baseline.

table 2 *Ten most frequently mentioned diseases in the patients' medical history by study arms (information taken from the general practitioner's information system using ICPC-codes*)*

DGIP (n = 61)		Usual care (n = 49)	
Disease – n (%)		Disease – n (%)	
Cataract	14 (22.9)	Essential hypertension	16 (32.6)
Dementia	14 (22.9)	Cataract	15 (30.6)
Depression	13 (21)	Diabetes	13 (26.5)
Essential hypertension	10 (16.3)	Dementia	12 (24.4)
Diabetes	10 (16.3)	Presbyacusis	8 (16.3)
Stroke	10 (16.3)	Atrial fibrillation	7 (14.2)
Atrial fibrillation	9 (14.7)	Stroke	7 (14.2)
Heart failure	9 (14.7)	Cholecystitis/cholelithiasis	6 (12.2)
Other diseases urinary tract	8 (13.1)	Depression	6 (12.2)
Diverticula/diverticulitis	7 (11.4)	Macular degeneration	5 (10.2)

*International Classification of Primary Care[31]

table 3 *The effect of the intervention on caregiver burden and time spent on caring for the patient at three and six months follow-up**

	3 months follow-up (T_1)			6 months follow-up (T_2)		
Outcome measure	D	CI	N	D	CI	N
Zarit Burden Interview†	-0.67	-4.03 – 2.69	92	2.29	-1.61 – 6.19	88
Time spent on personal care (hours/week)	0.53	-3.23 – 4.29	80	0.97	-1.66 – 3.60	84
Time spent on other assistive activities (hours/week)	-0.14	-4.70 – 4.42	75	1.86	-2.23 – 5.95	81

* Intervention effect: difference (D) and its 95% Confidence Interval (CI) in changes over time from baseline to three and six months between intervention and control group, for the number of subjects (N) in analysis.
Model: *Outcome Measure Change from baseline* $= \beta_0 + \beta_1$*treatment arm* $+ \beta_2$*baseline value Outcome Measure* $+ \beta_3$*minimisation factors + random intercept physician*
† Primary outcome caregiver

table 4 *Caregiver burden and time spent on caring for the patient at three and six months follow-up: unadjusted outcome measure scores*

	3 months follow-up (T_1)		6 months follow-up (T_2)	
Outcome measure	Usual care	Home visits*	Usual care	Home visits*
Zarit Burden Interview, mean ± SD†	29.1 ± 14.0	28.0 ± 12.5	28.7 ± 13.6	30.7 ±12.5
Time spent on personal care (hours/week), median [Interquartile range]	0 [0 – 3.5]‡	0.5 [0 – 6.5]‡	1.0 [0 – 3.0]‡	0.7 [0 – 7.0]‡
Time spent on other assistive activities (hours/week), median [Interquartile range]	2.0 [1.0 – 7.0]‡	2.0 [0.5 – 8.5]‡	3.25 [1.0 – 6.0]‡	2.0 [0.5 – 7.0]‡

* Treatment under study
† Primary outcome caregiver
‡ For skewed variables median and quartiles are presented

The subgroup analyses showed that at six months the home visits programme improved burden with 6.2 points in caregivers living together, while burden worsened with 4.0 points in caregivers in the intervention group sharing no household compared to the control caregivers (p-value for this interaction 0.04) (Table 5). No other significant interactions were found.

table 5 *Subgroup analyses of the effect of the intervention on caregiver burden (ZBI) in association with possible effect modifiers**

	3 months follow-up (T_1)		6 months follow-up (T_2)	
Effect modifier	D	P interaction	D	P interaction
Patient MiniMental State Examination < 21 (n=42) *vs*	0.0	0.75	1.7	0.75
Patient MiniMental State Examination ≥ 21 (n=68)	-1.1		3.0	
Patient independent (n=94) *vs*	-0.6	0.42	2.3	0.19
Patient in home for the aged (n=15)	-5.2		-9.5	
Patient and caregiver live together (n=23) *vs*	-2.7	0.53	-6.2	0.04
Patient and caregiver live separately (n=85)	0.0		4.0	

* Intervention effect: difference (D) in change over time in Zarit Burden Interview from baseline to three and six months between intervention and control group for each of the strata of the possible effect modifiers and the p-value for the interaction (P interaction)

Discussion

Caregiver burden was high in the caregivers that participated in this study compared to reference populations [24, 25]. The high burden scores remained roughly constant over time. The nurse-led intervention model for frail older people that was tested in this randomised controlled trial had no positive effect on caregiver burden and the time the caregiver spent on caring.

Our initial hypothesis, based on previous studies, was that the intervention would have a beneficial effect on caregiver burden, despite the fact that our home visits programme was an intervention that aimed at the patient in the first place, not at the caregivers. Overall, the intervention did not show an improvement of caregiver burden. However, the interaction analyses showed that estimating an effect for all caregivers combined probably does not satisfactorily describe the effect of the intervention on caregiver burden (Table 5). The interaction between home visits and living arrangement modified the effect of the intervention. The intervention performed convincingly better in caregivers living together with the patient than in caregivers living separately. This observation certainly makes sense, because the involvement of the caregivers also depends on their living arrangement: burden at baseline was almost eight points higher in caregivers living together, and they spent more time on caring. Because their needs were higher, caregivers sharing a household probably were more sensitive to intervention. In caregivers not sharing their household with the patient the feeling that the intervention only implied extra tasks may have prevailed. The intervention may also have made them more conscious of the vulnerability and impairments of their relative; information they simple did not have or were able to "neglect" in advance. It turned out that caregivers sharing a household not only were more involved in the caregiving process, they also were almost always spouse, older, and more frequently men. These are all factors that might interact with the intervention independently of living arrangement. However, because these other possible explanatory factors coincided to a considerable extent with

the subgroups living together or not in our study, it is difficult to tell whether it was really living arrangement and not for example the caregiver's sex that interacted with therapy. However, post-hoc subgroup analyses were performed on these separate factors, and neither showed a significant interaction effect on its own (data not shown). Our finding is in line with the results of a randomised study on caregiver support in dementia care that showed that beneficial effects were to be expected in caregivers sharing a household with the demented patient [26]. Other studies have also shown that caregiver burden and caregiver involvement may depend on living arrangement [27, 28]. This interaction may also be an explanation for the – in contradiction to ours – positive results found in other studies. While our study included no more than 25% caregivers living together, other studies included numbers of around 40-50% [11, 13, 14]. In our study, the intervention effect was significantly better in the group of caregivers sharing a household with the patients.

We used a robust design to evaluate the effect of home visits on caregiver outcomes. Nevertheless, there are some methodological issues that deserve attention. Our study arms were different in size. Numerical imbalance between study arms harms statistical efficiency, but this only occurs when the imbalance is larger than a factor 2 [29]. Patient randomisation minimised the possibility of selection bias. The comparison of patients' baseline characteristics (Table 1) and the results of sensitivity analyses confirmed this statement. Patients in the usual care group were only slightly older. Therefore, we did perform a sensitivity analysis adjusting for age of patient, which showed approximately the same results as the uncorrected analysis (data not shown). Ten to 20% of the included caregivers were not included in the primary analysis of burden, because of missing values. A sensitivity analysis of the primary outcome assigning the mean of the other group to the missing values gave about the same results [30]. While most caregivers cooperated enthusiastically, there was a group we failed to motivate. There are two reasons for this reluctance: the intervention is aimed at the patient in the first place, and we used a liberal definition of a caregiver in this study, identifying persons with a rather loose contact already as caregiver.

Our trial results showed that the effect of geriatric evaluation and management in the community on caregiver burden probably depends on the living arrangement of caregiver and patient; the intervention performed convincingly better in caregivers living together with the patient – who suffer the highest burden – than in caregivers living separately. This difference is probably caused by different levels of involvement of the caregivers. This result deserves further attention in future research. With increasing popularity of primary care alternatives for vulnerable older people, attention for the caregiver is a necessity, both in developing health care interventions, as in their evaluation and implementation.

References

1. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. *JAMA* 2002 February 27;287(8):1022-8.
2. Melis RJF, Olde Rikkert MG, Parker SG, van Eijken MIJ. What is intermediate care? *BMJ* 2004 August 14;329(7462):360-1.
3. WHO. Active ageing: a policy framework. Geneva: World Health Organisation; 2002.
4. Salvage AV, Jones DA, Vetter NJ. Options of people aged over 75 years on private ad local authority residential care. *Age Ageing* 1989;18:380-6.
5. Vernooij-Dassen M, Dautzenberg M. Collaboration between lay and professional care. In: Jones R, Britten N, Culpepper L, Gass DA, Grol R, Mant D et al., editors. *Oxford textbook of primary care*.Oxford: Oxford University Press; 2003.
6. Opie J, Rosewarne R, O'Connor DW. The efficacy of psychosocial approaches to behaviour disorders in dementia: a systematic literature review. *Aust N Z J Psychiatry* 1999 December;33(6):789-99.
7. Elkan R, Kendrick D, Hewitt M et al. The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature. *Health Technol Assess* 2000;4(13):i-339.
8. Pinquart M, Sorensen S. Helping caregivers of persons with dementia: which interventions work and how large are their effects? *Int Psychogeriatr* 2006 December;18(4):577-95.
9. Sorensen S, Pinquart M, Duberstein P. How effective are interventions with caregivers? An updated meta-analysis. *Gerontologist* 2002 June;42(3):356-72.
10. Lyons KS, Zarit SH, Sayer AG, Whitlatch CJ. Caregiving as a dyadic process: perspectives from caregiver and receiver. *J Gerontol B Psychol Sci Soc Sci* 2002 May;57(3):195-204.
11. Weuve JL, Boult C, Morishita L. The effects of outpatient geriatric evaluation and management on caregiver burden. *Gerontologist* 2000 August;40(4):429-36.
12. Tibaldi V, Aimonino N, Ponzetto M et al. A randomized controlled trial of a home hospital intervention for frail elderly demented patients: behavioral disturbances and caregiver's stress. *Arch Gerontol Geriatr Suppl* 2004;(9):431-6.
13. Hung LC, Liu CC, Hung HC, Kuo HW. Effects of a nursing intervention program on disabled patients and their caregivers. *Arch Gerontol Geriatr* 2003 May;36(3):259-72.
14. Silverman M, Musa D, Martin DC, Lave JR, Adams J, Ricci EM. Evaluation of outpatient geriatric assessment: a randomized multi-site trial. *J Am Geriatr Soc* 1995 July;43(7):733-40.
15. Melis RJF, Adang E, Teerenstra S et al. The cost-effectiveness of a multidisciplinary intervention model for community-dwelling frail older people. J Gerontol A Biol Sci Med Sci. In press 2007.
16. Melis RJF, van Eijken MIJ, Borm GF et al. The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]. *BMC Health Serv Res* 2005 October 5;5(1):65.
17. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975 November;12(3):189-98.
18. 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 December;43(11):1601-10.
19. Kempen GI. The MOS Short-Form General Health Survey: single item vs multiple measures of health-related quality of life: some nuances. *Psychol Rep* 1992 April;70(2):608-10.
20. Borm GF, Melis RJF, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. *Stat Med* 2005 December 15;24(23):3535-47.
21. Teerenstra S, Melis RJF, Peer PG, Borm GF. Pseudo cluster randomization dealt with selection bias and contamination in clinical trials. *J Clin Epidemiol* 2006 April;59(4):381-6.
22. Vernooij-Dassen MJ, Felling AJ, Brummelkamp E, Dauzenberg MG, van den Bos GA, Grol R. Assessment of caregiver's competence in dealing with the burden of caregiving for a dementia patient: a Short Sense of Competence Questionnaire (SSCQ) suitable for clinical practice. *J Am Geriatr Soc* 1999 February;47(2):256-7.
23. Duijnstee M, Guldemond H, Hendriks L. *Zorgkompas voor mantelzorgers van ouderen en chronisch zieken. Leidraad voor het in kaart brengen van de belasting van zorgende familieleden [care compass for informal carers of elderly people and people with a chronic illness. A guide for mapping the burden of caring family members]*. 1 ed. Utrecht: Nederlands Instituut voor Zorg en Welzijn/NIZW; 2001.
24. O'Rourke N, Tuokko HA. Psychometric properties of an abridged version of The Zarit Burden Interview within a representative Canadian caregiver sample. *Gerontologist* 2003 February;43(1):121-7.
25. Sambrook R, Herrmann N, Hebert R et al. Canadian Outcomes Study in Dementia: study methods and patient characteristics. *Can J Psychiatry* 2004 July;49(7):417-27.
26. Vernooij-Dassen M, Lamers C, Bor J, Felling A, Grol R. Prognostic factors of effectiveness of a support program for caregivers of dementia patients. *Int J Aging Hum Dev* 2000;51(4):259-74.
27. Deimling GT, Bass DM, Townsend AL, Noelker LS. Care-related stress. A comparison of spouse and adult-child caregivers in shared and separate households. *J Aging Health* 1989 February;1(1):67-82.
28. Ebly EM, Hogan DB, Rockwood K. Living alone with dementia. *Dement Geriatr Cogn Disord* 1999 November;10(6):541-8.
29. Pocock SJ. *Clinical trials : a practical approach*. Chichester: 1983.
30. Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med* 2001 December 30;20(24):3931-46.
31. Lamberts H, Wood M. *ICPC. International Classification of Primary Care*. Oxford: Oxford University Press; 1987.

Chapter 8

The cost-effectiveness of a multidisciplinary intervention model for community-dwelling frail older people

R.J.F. Melis, E. Adang, S. Teerenstra, M.I.J. van Eijken, A. Wimo,
T. van Achterberg, E.H. van de Lisdonk, M.G.M. Olde Rikkert

J Gerontol A Biol Sci Med Sci, in press (revised)

Objective: There is growing interest in geriatric care for community-dwelling older people. There are, however, relatively few reports on the economics of this type of care. This paper reports about the cost-effectiveness of the Dutch Geriatric Intervention Programme (DGIP) compared to usual care in frail older people at six months follow-up from a health care system's point of view.

Methods: We conducted this economic evaluation in an observer blind randomised controlled trial (Dutch EASYcare Study: ClinicalTrials.gov Identifier NCT00105378). Difference in treatment effect was calculated as the difference in proportions of successfully treated patients (prevented functional decline accompanied by improved well-being). Incremental treatment costs were calculated as the difference in mean total care costs. The incremental cost-effectiveness ratio (ICER) was expressed as total costs per successful treatment. Bootstrap methods were used to determine confidence intervals for these measures.

Results: The average cost of the intervention under study (DGIP) was 998 euros [95% CI 888 – 1108]. The increment in total costs resulting from DGIP was little over 761 euros [-3336 – 4687]. Hospitalisation and institutionalisation costs were less; home care, adult day care, and meals-on-wheels costs were higher. There was a significant difference in proportions of successful treatments of 22.3% [4.3 – 41.4]. The number-needed-to-treat was approximately 4.7 [2.3 – 18.0]. The incremental cost-effectiveness ratio is 3418 euro per successful treatment [-21458 – 45362]. The new treatment is cost-effective at a willingness-to-pay of 34000 euros.

Conclusions: The results of this economic evaluation suggest that DGIP is an effective addition to primary care for frail older people at a reasonable cost.

Introduction

There is growing interest in geriatric care for community-dwelling older people [1, 2]. Both growing awareness of patient autonomy and population ageing drive this development [3, 4]. Community-based geriatric intervention models aim to improve functioning, well-being, coping styles, and mortality among community-dwelling older persons. The idea is that increasing older people's autonomy will also increase their quality of life. Despite some promising effects on disabilities, mood, and mortality, there is still much controversy about important determinants of success such as which patients to include, the setting and the intensity of intervention [1, 5, 6]. However, for geriatric primary care to be truly valuable to society it needs to provide effective interventions that give 'value for money'. There are relatively few reports on the economics of outpatient comprehensive geriatric assessment or preventive home visits [7, 8]. A review on comprehensive geriatric assessment concluded that evidence suggests no increase of total costs of care [7]. Another review about preventive home visits stated that there is a potential for produce net cost savings [8].

From April 2003 until July 2005 we carried out a randomised comparison of a multidisciplinary home-based intervention programme for frail older people with usual care (Dutch EASYcare Study) [9]. It showed that patients' well being and functional abilities improved compared to control conditions [10]. Subjective caregiver burden seemed to improve in caregivers who shared a household [11]. This paper reports about health care utilisation and the cost-effectiveness of the Dutch Geriatric Intervention Programme (DGIP) compared to usual care in frail older people from a health care system's point of view.

Methods

Study design and setting

We conducted an economic evaluation in an observer blind randomised controlled trial (ClinicalTrials.gov Identifier NCT00105378) of which the design was previously published [9]. We used a health care system's perspective to identify all relevant costs. The local ethical committee approved of this study.

Patients

General practitioners referred frail older patients to our model when there was a problem in cognition, nutrition, behaviour, mood, or mobility. The problem had to call for nursing assessment, coordination of care, or therapeutic monitoring and case management. Patients were living in their own home or in a home for the aged and were 70 years or older. We restricted patient inclusion to those with limitations in cognition, (instrumental) activities of daily life, or mental well-being.

Randomisation

Patients were randomised over two treatments arms: DGIP and usual care. We used a two-step pseudo cluster randomisation procedure, because we expected methodological problems when using either individual randomisation or cluster randomisation [9, 12-14].

Interventions

A geriatric specialist nurse visited the patients randomised to the intervention arm. During a maximum of three months up to six visits for additional geriatric evaluation and management were performed. The nurse, geriatrician, and general practitioner had frequent consultations on individual cases. No restrictions were imposed on the care patients received in the usual care group. However, no model of care comparable to DGIP was regularly available in usual care.

Data collection

Effect measures

The patients were interviewed to obtain written informed consent and to collect baseline demographic characteristics, data on general health conditions, and data on use of home care. These measurements were repeated six months after inclusion. The outcome measures used for this economic evaluation were functional performance in (instrumental) activities of daily living measured using the Groningen Activity Restriction Scale and mental well-being using the subscale mental health of MOS-20.[15, 16]

Cost analysis

Costs were measured by registration of volumes of care consumed and multiplied by appropriate prices per unit of care. To be able to calculate the costs of DGIP, DGIP-nurses registered the time spent on the intervention using an MS Outlook® agenda. They registered the visits per patient, and the time spent on consultation, phone calls, travelling, and administration. Data on the workload of the general practitioner and the geriatrician resulting from DGIP were derived from the workload of the nurses. We also established the amount of care consumed during the follow-up period of six months. In order to calculate incremental costs, data collection of cost measures focused on cost

drivers, i.e. these factors that were likely to cause the most important cost differences between strategies. These were: utilisation of family physician care (number of practice visits, house calls, and telephone calls), the number of referrals to other health care providers (outpatient specialist care, paramedical disciplines such as physical therapist, occupational therapist etc.), and the number of days in hospital. All the above information was taken from the general practitioner's information system. The number of days hospitalised was crosschecked against the information provided by the patient using questionnaires at three and six months of follow-up. The patients provided information on the amount of home care, adult day care, and meals-on-wheels they used during follow-up. They also provided information on number of days institutionalised in nursing home or home for the aged; this information was crosschecked against the information available from the general practitioner's information system.

Cost prices per unit of care (Euros) were taken from Dutch guidelines for economic health care evaluations [17]. In accordance with these guidelines, 45% overhead costs were added to the total direct costs to calculate a cost price per DGIP visit. Cost prices of referrals to outpatient specialist care, and paramedical disciplines were estimated because this information was not readily available. All prices were indexed to the price level of 2005, using the Dutch consumer price index figures for health care costs[18].

table 1 *Demographic characteristics and outcome measures of the study population at enrolment**

	Usual care (N = 66)	Dutch Geriatric Intervention Programme (N = 85)	P
Age – yr	82.8±6.6	81.7±5.9	0.22
Female sex – no. (%)	49 (74.2)	64 (75.3)	0.99
Marital status – no. (%)			0.31
• Married	19 (29.2)	23 (27.4)	
• Divorced	2 (3.1)	5 (6.0)	
• Widow(er), partner deceased	36 (55.4)	47 (56.0)	
• Single	8 (12.3)	8 (9.5)	
• Living together unmarried	0	1 (1.1)	
Living in home for the aged – no. (%)	11 (16.9)	12 (14.1)	0.56
Informal caregiver – no. (%)	49 (74.2)	65 (76.5)	0.71
Receiving home care – no. (%)	34 (51.5)	43 (51.1)	0.97
Groningen Activity Restriction Scale-3	34.1±8.7	35.3±8.1	0.68
Medical Outcomes Scale-20 mental health	53.8±17.7	53.3±20.9	0.87
MOS-20 physical functioning	16.7 [0 – 33.3]§	16.7 [0 – 36.7]§	1.00
MOS-20 role functioning	0 [0 – 50.0]§	0 [0 – 50.0]§	0.87
Cumulative Illness Rating Scale-Geriatrics	9.8±4.3	10.2±3.7	0.64
Timed Up and Go – seconds	16.0 [12.0 – 22.0]§	16.0 [12.0 – 25.0]§	0.60
MMSE	22.0±6.0	22.8±5.5	0.22
Cantril's self anchoring ladder	5.9±2.1	5.7±2.1	0.78
Dementia Quality of Life positive affect	3.1±0.7	3.2±0.7	0.50
DQoL negative affect	2.6±0.6	2.7±0.7	0.24
Loneliness Scale De Jong-Gierveld	5.7±3.6	4.4±3.2	0.04

* For all – except skewed – variables means ± SDs are presented, for skewed variables (§) median and interquartile range are presented.

Statistical analysis

To obtain an interpretable effect measure, we considered the treatment a success if a patient's MOS-20MH score increased by more than 10 points and GARS-3 score declined no more than 4.5 points. We used this definition because the course of GARS-3 scores showed a decline in our control group, while fairly steady MOS-20MH scores were observed over a six months period. Relative to the score range, both cut-offs represent approximately 10% change in score.

Difference in treatment effect was calculated as the difference in proportions of successfully treated patients. Incremental treatment costs were calculated as the difference in mean total care costs. The incremental cost-effectiveness ratio (ICER) was expressed as total costs per successful treatment. Bootstrap methods were used to determine confidence intervals for these treatment arms' differences that take into account the clustering of patients within their general practitioner. From the bootstrap sample a cost-effectiveness acceptability curve (CEAC) was drawn. This curve gives an estimate of the proportion of the bootstrap distribution favouring one strategy over the other given a willing-to-pay (WTP) for a gained unit of effect 19. We performed a sensitivity analysis for missing data; missing values were substituted with the mean of the other group 20.

Results

In this study 151 patients were included. Our study population consisted in majority of widowed females born in the Netherlands. A majority (85%) lived in their own home; the rest lived in homes for the aged. The patients had a mean age of 82.2 years, a lot of co-morbidity, considerable cognitive deterioration, and low mental well-being. Approximately half the study group had home care available at baseline.

table 2 *Number of health care units used per patient during six months follow-up*

	Usual care (n = 66) Mean ± SD	DGIP* (n = 85) Mean ± SD
DGIP visits	0	3.8 ± 1.3
Hospitalisation (days)	6 ± 16	4 ± 14
Physician care (number)	10.7 ± 8.9	13.4 ± 8.2
• Practice visits	1.0 ± 1.9	1.0 ± 1.8
• House calls	2.2 ± 3.3	2.6 ± 3.5
• Telephone contacts	1.5 ± 2.3	2.4 ± 2.9
• Administrative contacts	3.4 ± 3.9	4.6 ± 3.6
• Emergency practice visits	0	0.1 ± 0.2
• Emergency house calls	0.5 ± 0.6	0.4 ± 0.7
• Emergency tel. contacts	0.2 ± 0.4	0.7 ± 0.9
• Unspecified	2.4 ± 3.3	2.5 ± 2.9
• Unspecified, emergency	0.1 ± 0.3	0.1 ± 0.3
Institutionalisation (days)		
• Home for the aged	32 ± 65	24 ± 58
• Nursing home	5 ± 23	4 ± 16
Home care (hours)	63.4 ± 77.0	88.6 ± 172.4
• Domestic care	38.7 ± 44.5	33.3 ± 39.0
• Personal care	22.7 ± 43.0	52.4 ± 164.5
• Nursing care	2.0 ± 9.3	2.9 ± 11.5
Day care (days)	3 ± 10	6 ± 21
Meals-on-wheels (days)	33 ± 63	44 ± 66

* Dutch Geriatric Intervention Programme

table 3 *Treatment effect and health care usage and costs per patient at six months of follow-up*

	Usual care (n = 66) Mean ± SD	DGIP* (n = 85) Mean ± SD	Difference [95% CI]
EFFECT			
GARS-3†‡	-1.6 [-3.9 – 0.7]		
MOS-20MH†§	9.1 [2.4 – 15.6]		
"Successful treatments" (%)	7 (13.7)	27 (36.0)	22.3 [4.3 – 41.4]
Number of missing values	15	10	
Number needed to treat	4.7 [2.3 – 18.0]		
COSTS||			
Total care costs (euros)	8952 ± 9757	9713 ± 10205	761 [-3336 – 4687]
Intervention (DGIP)	0	998 ± 408	998 [888 – 1108]
Hospitalisation	2291 ± 6449	1616 ± 5480	-675 [-3275 – 1375]
• Number (%)¶	14 (21.2)	11 (12.9)	-8.3 [26.4 – 7.2]
• Costs if used**	10801 ± 10437	12487 ± 10147	1687 [8637 – 11592]
Physician care	239 ± 207	277 ± 209	38 [-43 – 117]
• Number (%)¶	65 (98.5)	83 (97.9)	-0.8 [-6.3 – 5.4]
Outpatient care	93 ± 264	71 ± 207	-21 [-130 – 68]
• Number (%)¶	16 (24.2)	22 (26.9)	1.6 [-16.2 – 18.3]
• Cost if used**	382 ± 429	275 ± 335	-107 [-453 – 180]
Paramedical care	43 ± 167	65 ± 156	22 [-58 – 78]
• Number (%)¶	5 (7.6)	15 (17.7)	10.1 [-3.9 – 21.7]
• Cost if used**	567 ± 289	368 ± 164	-199 [-578 – 93]
Institutionalisation	3727 ± 7279	2886 ± 5802	-841 [-3586 – 1655]
• Number (%)¶	17 (25.8)	19 (22.4)	-3.4 [-20.2 – 13.5]
• Cost if used**	14470 ± 7077	12910 ± 4527	-1560 [-6274 – 3197]
Home care	1880 ± 2483	2832 ± 5887	952 [-542 – 2799]
• Number (%)¶	40 (60.6)	61 (71.8)	11.1 [-7.6 – 29.5]
• Cost if used**	3101 ± 2528	3946 ± 6637	845 [-1106 – 3241]
Day care	380 ± 1084	622 ± 2282	241 [-294 – 1053]
• Number (%)¶	10 (15.2)	10 (11.8)	-3.4 [-17.1 – 11.4]
• Cost if used**	2511 ± 1596	5383 ± 4607	2772 [-420 – 7028]
Meals-on-wheels	299 ± 560	390 ± 595	91 [-150 – 378]
• Number (%)¶	18 (27.3)	34 (40.0)	12.7 [-6.9 – 34.6]
• Cost if used**	1098 ± 521	976 ± 560	-122 [-475 – 329]

* Dutch Geriatric Intervention Programme

† Difference between study arms in outcome measures' change from baseline at 6 months follow-up

‡ Groningen Activity Restriction Scale (GARS-3) runs from 18 to 54, with 18 indicating best score

§ Medical Outcomes Scale 20 subscale mental health (MOS-20MH runs from 0 to 100, with 100 indicating best score

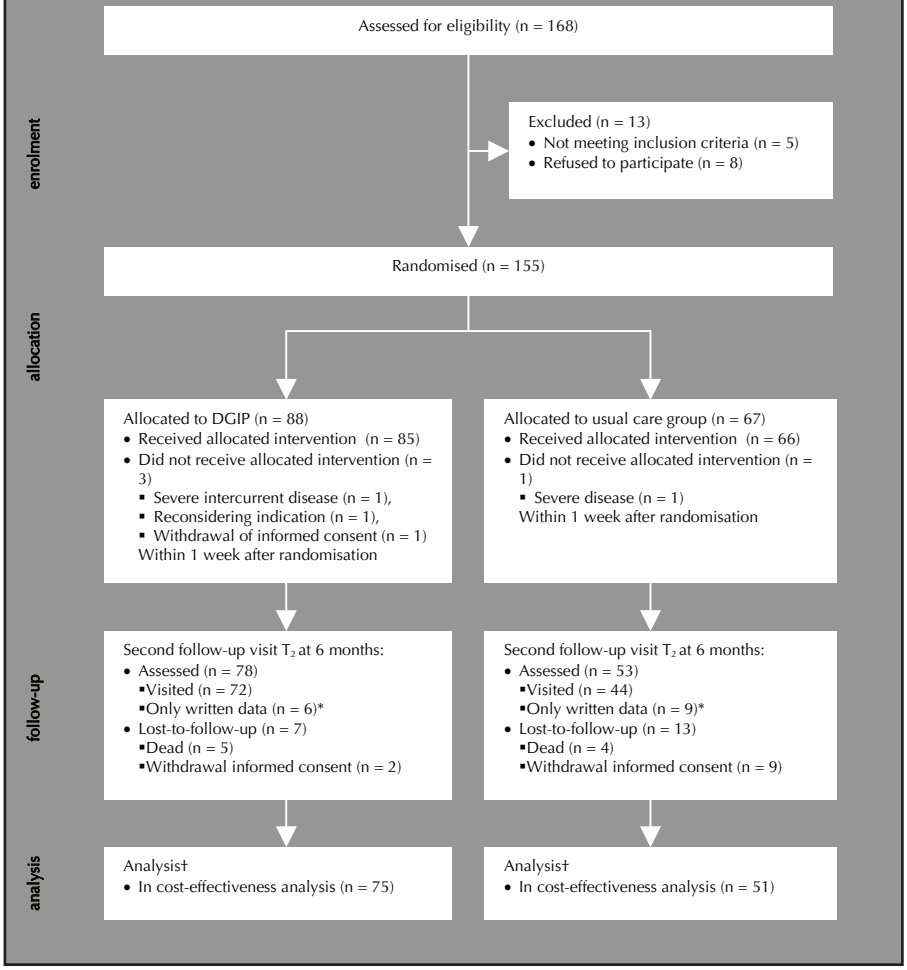
|| Costs per patient were averaged over all patients in each group

¶ Number: number of patients who used a specific service

** Cost if used: costs of a specific service when averaged over these patients who actually used a certain service

In the group that received DGIP 85 patients were included, and 66 were included in the usual care group. Baseline characteristics and baseline measures of primary outcomes showed no significant differences between study groups, except for the score on loneliness (table 1). Figure 1 shows the participant flow through the study. The nurse visited DGIP patients almost four times (table 2). Of all patients 34 patients were successfully treated: they experienced an improvement in well-being without a decline in functional performance (table 3). Of these patients 7 were in the control group, and 27 in the intervention group. The significant difference in proportions of successfully treated patients was 22.3% [4.3 – 41.4] in favour of DGIP, and the number needed to

figure 1 *Study flow chart*



* A number of participants withdrew their consent for the visit for data acquisition by interview, but agreed to fill in the questionnaires.

† Differences between numbers assessed at follow-up and numbers included in the analysis result from the fact that although assessment was executed, participants not always completely filled the instruments. Therefore, responder status could not be established for everyone assessed.

treat was 4.7 [2.3 – 18.0]. The DGIP intervention cost 998 euros per patient [95% CI 888 – 1108]. The incremental cost of the nursing programme was 761 euros [-3336 – 4687]. Hospitalisation (-675 euros) and institutionalisation in home for the aged and nursing home (-841 euros) cost less in the DGIP group, while home care (+952 euros), day care (+241 euros) and meals-on-wheels (+91 euros) were more expensive (table 3). The incremental cost-effectiveness ratio is 3418 euros per successful treatment [-21458 – 45362]. The probability of DGIP being the "dominant" intervention was estimated to be 34.6% (figure 2: proportion of incremental cost-effectiveness ratios falling in the Southeast quadrant). The vertical axis of the CEAC represents the probability that the incremental cost-effectiveness ratio of DGIP compared with usual care is acceptable for

a range of values of the willingness to pay per successful treatment (figure 3). If society is willing to pay 34000 euros or more for a successful treatment than there is a 95% probability that DGIP is efficient. During follow-up 25 patients had missing values for treatment success; 10 patients were in the intervention group, 15 in the control group (figure 1). However, the results from the sensitivity analysis for missing data did not essentially differ from the results presented (data not shown).

Discussion

In this group of frail patients the mean total care costs in the control group were approximately 9000 euros over a six months period. The yearly health care costs in the age group 75-84 year is 8408 euros in the Netherlands [21], but a vulnerable population – in which higher health care costs can be expected – was included in this study. The average cost of the intervention under study (DGIP) was almost 1000 euros. The increment in total costs resulting from DGIP was little over 750 euros. Hospitalisation and institutionalisation costs were less; home care, adult day care, and meals-on-wheels costs were higher. There was a difference in proportions of successful treatments (prevented functional decline accompanied by improved well-being) of about 20%. The number-needed-treat was approximately 5. The cost-effectiveness ratio is roughly 3500 euros per successful treatment. The new treatment is cost-effective at a willingness-to-pay of 34000 euros.

Comparison of these results with literature is difficult, because there are not many cost-effectiveness studies in this field. Most papers only provide data on incremental costs without calculating cost-effectiveness ratios [7]. The available reviews of literature hint at the possibilities of cost savings, but request more research into this matter at the same time [7, 8]. One paper on preventive home visits concluded that the visits produced net cost savings in the third year [22]. However, no information on the uncertainty for this estimation was given, which limits interpretation. A sound cost-effectiveness evaluation of outpatient geriatric evaluation and management estimated a cost-effectiveness ratio of 10600 dollars per QALY, and 26500 per QALY if the time horizon of five years was limited to the follow-up period of 64 weeks [23]. Another randomised study of outpatient geriatric evaluation and management concluded there were no increases in costs [24]. Our study did show a small increment in costs, although the cost-effectiveness seems reasonable. The uncertainty in the estimation is large, because the differences in total care costs are large. Total care costs were less than 50 euros for some patients, for others the amount was larger than 40000 euros. This reflects the heterogeneity in the care needed.

An explanation for the positive results in some of the earlier studies might be that these studies focused on certain aspects of health care, such as hospital costs [7]. Our study took a wide health care system's perspective as starting point. Our study showed lower hospital and institutionalisation costs, but home care and adult day care costs were higher. This is probably a direct intervention effect, because often nurses organised more home and adult day care. As independent living is considered an important aim [3], these increases seems reasonable and justifiable.

We focused on all cost drivers relevant for the viewpoint taken. Medication, laboratory, and diagnostics costs were not taken into account, because the influence of DGIP on these

figure 2 *Cost-effectiveness plane for the incremental costs and effects of DGIP compared to usual care*

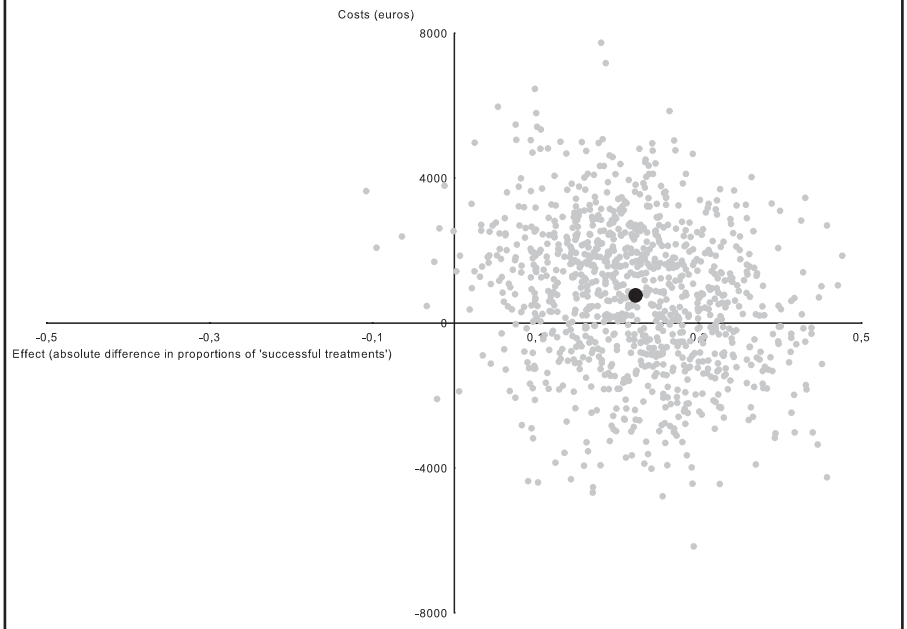
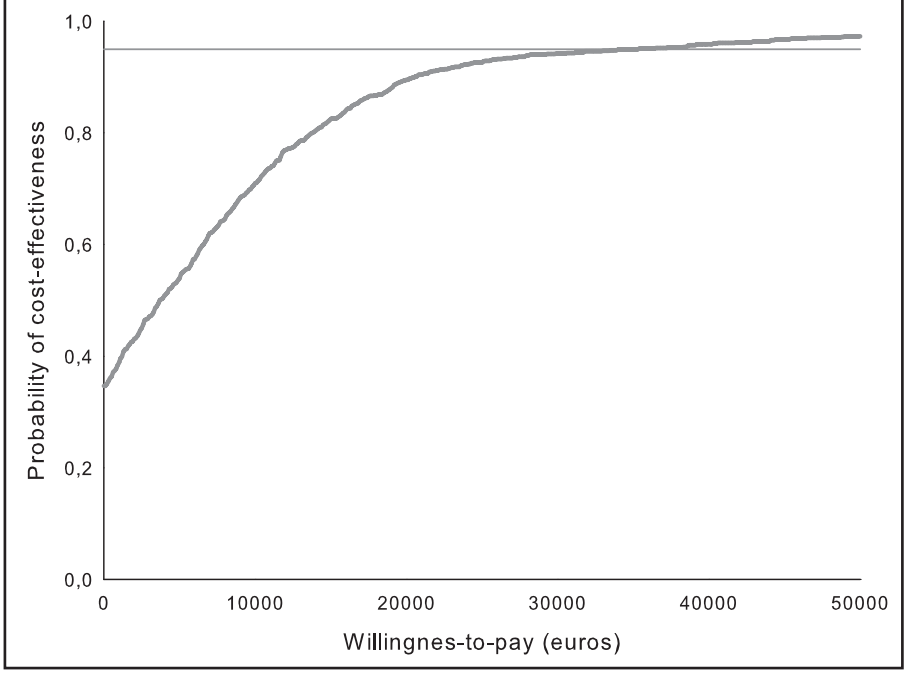


figure 3 *Cost-effectiveness acceptability curve of DGIP compared to usual care*



cost parameters was judged to be negligible. For the same reason no societal perspective was used. Given their age, the effect of the intervention on patients' productivity is very limited. The effect on caregivers' productivity is probably fairly small as well; caregiver burden and time spent on care was hardly affected by the intervention.

No QALY's were used in our study, which limits comparability with other interventions. On the other hand, the use of QALY's in gerontological research is delicate and its interpretation difficult. Generic instruments like EuroQoL from which QALY's can be calculated probably show floor effects in frail patients as in our study. Also health states of frail older people are complex and often not sufficiently graded in terms of QALY's.

Of course, every operationalisation of "successful treatment" is arbitrary to a certain degree. However, the difference in successfully treated patients, which was 22.3% in favor of DGIP with the currently used dichotomisation, is rather insensitive to change of cut-offs. Varying the cut-offs on GARS-3 and MOS-20MH in different combinations, resulted in differences in treatment success rates roughly ranging from 15 to 25% all in favour of DGIP (data not shown).

This economic evaluation was conducted alongside a randomised trial, which is an important safeguard against threats to internal validity. Study arm sizes were somewhat different, but patients were very comparable at baseline. The number of missing values (20%) was expected given their frailty. Participants who were lost-to-follow-up were older and had worse GARS-3, MOS-20MH, and MMSE scores at baseline. Moreover, when we used a conservative imputation strategy – assigning "mean of the other group" to missing patients, a sensitivity analysis showed about the same results [20].

External validity benefited from the wide health care system perspective this evaluation used. The study conditions were very similar to current practice; as in usual care many physicians were involved in referring patients. Also, no strict treatment protocols were used; the control group received usual medical care without any restrictions, and guidelines led the DGIP intervention without restraining free selection from the available treatment options.

The results of this economic evaluation suggest that DGIP is an effective addition to primary care for frail older people at a reasonable cost. However, there is a need for larger and more rigorously designed economic evaluations of this type of health care models, because uncertainty in the cost-effectiveness of these models is still large.

References

1. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. *JAMA* 2002 February 27;287(8):1022-8.
2. Melis RJF, Olde Rikkert MG, Parker SG, van Eijken MIJ. What is intermediate care? *BMJ* 2004 August 14;329(7462):360-1.
3. WHO. Active ageing: a policy framework. Geneva: World Health Organisation; 2002.
4. Salvage AV, Jones DA, Vetter NJ. Options of people aged over 75 years on private ad local authority residential care. *Age Ageing* 1989;18:380-6.
5. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993 October 23;342(8878):1032-6.
6. Fletcher AE, Price GM, Ng ES et al. Population-based multidimensional assessment of older people in UK general practice: a cluster-randomised factorial trial. *Lancet* 2004 November 6;364(9446):1667-77.
7. Wieland D. The effectiveness and costs of comprehensive geriatric evaluation and management. *Crit Rev Oncol Hematol* 2003 November;48(2):227-37.
8. Elkan R, Kendrick D, Hewitt M et al. The effectiveness of domiciliary health visiting: a systematic review of international studies and a selective review of the British literature. *Health Technol Assess* 2000;4(13):i-339.
9. Melis RJF, van Eijken MIJ, Borm GF et al. The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]. *BMC Health Serv Res* 2005 October 5;5(1):65.
10. Melis RJF, van Eijken MI, Teerenstra S et al.; A randomised study of a multidisciplinary program to intervene on geriatric syndromes in frail older people who live at home (Dutch EASYcare Study) [Clinicaltrials.gov Identifier NCT00105378]. *J Gerontol A Biol Sci Med Sci*;in press.
11. Melis RJF, van Eijken MI, van Achterberg T, Vernooij-Dassen M, van de Lisdonk EH, Olde-Rikkert MG; The effect on caregiver burden of a problem-based home visiting program for frail older people in a randomized study. submitted.
12. Teerenstra S, Melis RJ, Peer PG, Borm GF. Pseudo cluster randomization dealt with selection bias and contamination in clinical trials. *J Clin Epidemiol* 2006 April;59(4):381-6.
13. Teerenstra S, Melis RJF, Peer PG, Borm GF. Pseudo cluster randomization deals with selection bias and contamination in clinical trials. *J Clin Epidemiol* 2005;in press.
14. Borm GF, Melis RJF, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. *Stat Med* 2005 December 15;24(23):3535-47.
15. 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 December;43(11):1601-10.
16. Kempen GI. Het meten van de gezondheidstoestand van ouderen. Een toepassing van een Nederlandse versie van de MOS-schaal. [Assessment of health status of the elderly. Application of a Dutch version of the MOS scale]. *Tijdschr Gerontol Geriatr* 1992 August;23(4):132-40.
17. Oostenbrink JB, Bouwmans CAM, Koopmanschap MA, Rutten FFH. Handleiding voor kostenonderzoek, methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. 2004. Diemen, College voor zorgverzekeringen.
18. StatLine. Voorburg: Statistics Netherlands; 2005.
19. Lothgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. *Health Econ* 2000 October;9(7):623-30.
20. Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med* 2001 December 30;20(24):3931-46.
21. Polder JJ, Takken J, Meerding WJ, Kommer GJ, Stokx LJ. kosten van ziekten in Nederland. De zorgeuro ontrafeld. 2002. Bilthoven, Rijksinstituut voor Volksgezondheid en Milieu.
22. Stuck AE, Minder CE, Peter-Wuest I et al. A randomized trial of in-home visits for disability prevention in community-dwelling older people at low and high risk for nursing home admission. *Arch Intern Med* 2000 April 10;160(7):977-86.
23. Keeler EB, Robalino DA, Frank JC, Hirsch SH, Maly RC, Reuben DB. Cost-effectiveness of outpatient geriatric assessment with an intervention to increase adherence. *Med Care* 1999 December;37(12):1199-206.
24. Cohen HJ, Feussner JR, Weinberger M et al. A controlled trial of inpatient and outpatient geriatric evaluation and management. *N Engl J Med* 2002 March 21;346(12):905-12.

Chapter 9

Discussion

Principal findings

This thesis aimed to study the effects of the multidimensional Dutch Geriatric Intervention Programme (DGIP), consisting of a home visiting programme for vulnerable older people living in their own home. To enable us to study these objectives in a valid way, a methodological innovation of the randomisation procedure was deemed necessary. The second aim of this thesis was therefore to describe and evaluate the method of pseudo cluster randomisation.

This study followed 151 patients and 110 caregivers for six months. 85 patients (and 61 caregivers) were randomly included in the intervention programme, and 66 patients (and 49 caregivers) received usual care. Patients were very vulnerable, and caregivers highly burdened.

The new intervention model showed positive effects on self-reported functional abilities and mental well-being of vulnerable older people. These positive effects persisted after discontinuing the home visits. This problem-based care model might have implied some extra burden for caregivers living separately. On the other hand, in caregivers sharing a household the intervention seemed to improve caregiver burden. Possibly, this difference is explained by baseline differences in perceived burden of both subgroups: the intervention decreased caregiver burden in the most burdened group of caregivers. Problem-based nursing care can effectively support primary health care for vulnerable older people, and does so at reasonable cost.

Methodological aspects of the Dutch EASYcare Study

Despite the careful design of this study, several methodological aspects deserve consideration; how did these aspects influence the validity and generalisability of the results? The most important issues are the validity of the application of pseudo cluster randomisation, and the effect of attrition.

Studying frail older people:

missing values, and unblinding the outcome assessment

We aimed to study the effect of this intervention in frail older people, because – for reasons pointed out in detail in chapter 1 – it is for this group in particular that new primary care strategies have to be developed for [1]. The age, co-morbidity, and GARS-3 scores at baseline and the overall mortality during follow-up showed that in this study a group of very old, and vulnerable patients was indeed sampled. The CIRS-G score was comparable to the score inpatients had at admission to the geriatric ward of our hospital [2]. Studying a frail older population has proven to be difficult, and specific procedures are needed to ensure the conduct of a valid and useful study [1]. Several precautions were taken in the Dutch EASYcare Study: we accounted for attrition in the sample size considerations, visited the patients at home for baseline- and outcome assessment, included a specific consent procedure for patients with impaired decision-making capacities, balanced the number, intensity, and length of the assessment visits and questionnaires, and defined a limited data set for patients for whom the complete assessment turned out to be too demanding.

As could be expected, attrition occurred. Drop outs occurred mainly because patients (or their caregivers) felt participation in follow-up visits for effect measurement was too burdensome while it provided no further benefit. No patients declined follow-up because of a reason that related to the intervention that was studied. Another more disguised mechanism of attrition occurred when patients were willing to participate, but were unable to complete the assessment instruments for instance because of cognitive impairments.

To study the effect of attrition on the outcomes, we performed a sensitivity analysis assigning the "mean of the other group" to the missing values, which is a conservative method [3]. "Replacing missing data by the mean of the other group has very good properties in regard of the type I error, the losses of power are acceptable if the drop-out rates are moderate and the drop-outs are not completely equal in the groups" [3]. A drawback is that imputing the mean of the competing group seems rather irrational from a clinical point of view. However, the use of "last-observation-carried-forward" was not an option either, because of the trend of decline which many outcomes show in an ageing population [3]. Multiple imputation is a very popular technique, but it is important to remember that most techniques used for multiple imputation assume that missing values are "missing at random" [4]. In other words, missing values are replaced with values of other non-missing subjects of the same group. As such, multiple imputation ignores an important source of potential bias: what was the reason for the subjects to drop out?

This trial was observer blind, because double blinding was impossible given the type of intervention under study. In a fairly high number of cases treatment assignment was revealed to the assessor during follow-up visits. Two strategies were tested in a pilot study phase: 1. the assessors explicitly asked the participants not to mention to the assessor whether a nurse visited them or not, and 2. the assessor avoided the subject of the house calls. The first approach did not work well, because participants often did not quite understand what they were asked. Having touched upon the subject of the house calls, participants often revealed their treatment assignments. One participant responded: "you mean I am not allowed to tell you that I was visited by nurse A?" This again shows how complicated research can be in frail subjects, who are not easily instructed. We decided to use the second strategy, but this was not very successful either because treatment assignment was for instance revealed through subtle hints such as a programme for day structure lying on the table. Maybe we should accept that an observer blind trial is impossible, if the outcome measures require the use of an assessor. Our primary outcome measures were taken using a written questionnaire which was filled out before each follow-up visit; the assessor has little influence on this.

Studying caregivers: difficult to engage

Between 10 and 20% of the included caregivers were not included in the primary analysis of burden, because of missing values. While most caregivers cooperated enthusiastically, there was a group we failed to motivate. There are two reasons for this reluctance: the intervention is aimed at the patient in the first place, and we used a liberal definition of a caregiver in this study, identifying persons with a rather loose contact already as caregiver. The definition we used for an informal caregiver was: a significant other person who felt responsible for the well-being of the patient. We used this rather liberal definition because informal caregivers are a very heterogeneous group of people and more strict definitions might fail to identify persons as caregivers when they actually feel they are.

Pseudo cluster randomisation:

assumptions justifying the use of pseudo cluster randomisation and the effect on contamination and study validity

Pseudo cluster randomisation was used in our trial because we assumed that randomisation at the level of the individual patient risked contamination bias, and that the accepted solution to avoid contamination bias, cluster randomisation, risked selection bias and recruitment problems. As these were the assumptions we had in advance, we considered the evidence we could find for these assumptions during the conduct of our trial. We found that general practitioners had strong preferences for randomisation of their patients to the intervention, and they also indicated that these are treatment preferences which influence their inclusion behaviour.

As randomisation occurred asymmetrically, general practitioners might get an idea about randomisation proportions, when recruitment in their own cluster advances. We found little evidence that general practitioners had this knowledge. Our general practitioners included only four patients on average, but even those with up to ten inclusions were not able to estimate accurately the randomisation ratio used. It seems that for cluster sizes of up to ten inclusions, asymmetrical randomisation in a 20:80 (= ratio intervention:control) or 80:20 does not lead to predictability of the randomisation sequence. We only had two general practitioners with more than ten patients in the study, therefore it is impossible to reach conclusions about the predictability of randomisation ratios with inclusion numbers over ten.

Unfortunately, the intervention group was somewhat larger than our control group. One might speculate whether this was a sign of some predictability of treatment assignment after all. The variation in the number of patients each physician included was very large. The two physicians who included more than 10 patients were assigned to the group of general practitioners whose patients were randomised to the intervention group in majority. This observation completely explains the unbalanced numbers of control and intervention group. One might not be satisfied and argue that these two general practitioners probably knew most of their patients were randomised to the intervention and be even more convinced about the occurrence of treatment predictability. However, the presence of these two physicians in the group where patients were randomised to the intervention group in majority, might just as well be a result of chance, as the number of often-recruiting physicians is so small. It would have been much more disturbing, if for example a majority out of ten or twenty often-recruiting general practitioners had been assigned to the group where patients were randomised to the intervention group in majority. Further reassurance is offered by the fact that patients in both study arms showed no signs of selection bias.

While we were able to evaluate the effect of pseudo cluster randomisation on the occurrence of selection bias with empirical information, we cannot judge whether contamination was effectively prevented or whether contamination was absent from the start. There are some indirect examples of the possible occurrence of contamination. One example is the fact that general practitioners indicated that they used elements from the intervention treatment when treating control group patients. Also the physicians' self-perceived competence in geriatrics increased. This might point to a learning effect, because it was not present in general practitioners who did not recruit.

Outcome measures:

how to define clinical relevance in vulnerable older people

We considered the treatment a success if a participant's MOS-20MH score increased by more than 10 points and GARS-3 score declined no more than 4.5 points. We used this definition because GARS-3 scores showed a decline in our control group, while fairly steady MOS-20MH scores were observed over a six months period. Of course, every responder definition is arbitrary to a certain degree. However, the difference in successfully-treated participants, which was 22.3% in favour of the new treatment using our dichotomisation, is rather insensitive to change of cut-offs. Varying the cut-offs on GARS-3 and MOS-20MH in different combinations, resulted in differences in treatment success rates roughly ranging from 15% to 25% all in favour of the new intervention treatment (table 1).

table 1 *Sensitivity analysis using different responder definitions in the responder analysis*

EFFECT Number of responders using different responder definitions (%)	Usual care (n = 66)	DGIP (n = 85)*	Difference [95% CI]
Change GARS < 2.25, MOS20MH ≥ 10	6 (11.8)	23 (30.7)	18.9 [5.2 – 32.6]
Change GARS < 4.5, MOS20MH ≥ 15	5 (9.8)	20 (26.7)	16.9 [4.0 – 29.8]
Change GARS < 4.5, MOS20MH ≥ 5	13 (25.5)	32 (43.2)	17.8 [1.3 -34.2]
Change GARS < 6.75, MOS20MH ≥ 10	7 (13.8)	29 (38.7)	24.9 [10.4 – 39.5]
Change GARS < 2.25, MOS20MH ≥ 5	12 (23.5)	27 (36.5)	13.0 [-3.0 – 29.0]
Change GARS < 6.75, MOS20MH ≥ 15	5 (9.8)	22 (29.3)	19.5 [6.4 – 32.7]

* Dutch Geriatric Intervention Programme

A final problem with health-related quality of life measures of the kind we used, is that they are criticised by patients with regard to their relevance in daily life [5]. Individualised, patient-centred goal setting based on patient preferences might provide a more sensitive and meaningful outcome measure [6]. In a trial conducted by Rockwood et al no effects were found using health-related quality of life measures [7], while goal attainment scaling showed effect [6]. Unfortunately, the application of these clinometric instruments in the practice of clinical trials can be difficult. In our study a comparison was made with usual care, a condition we wanted to be left unaffected by our study. However, the application of the instrument of goal attainment scaling requires the definition of individualised treatment goals, which is an intervention in itself.

Follow-up:

what is the optimal study length?

Our study applied a maximum follow-up of six months for the primary outcomes and up to two years (depending on whether patients were included early or late in the recruitment period) for survival, and institutionalisation. Six months of follow-up may be fairly short, but it is important to remember that there are many competing events in these vulnerable older people that cause a lot of "noise", this "noise" causes increasing variances in the outcome measures. Maybe a combination of short term as well as longer term (for instance 1 year) follow-up is useful. However, any study is a compromise between the most ideal study design and feasibility: much longer follow-up periods would not have been acceptable to the patients, or only at the expense of large numbers of dropouts.

Caregiver outcomes

Our main caregiver-related outcome measure was perceived caregiver burden. We also studied the (self-reported) effects on actual workload in terms of hours spent on caring, which was assessed by asking the caregivers to estimate the amount of time they spent in general on different tasks. These outcomes provide information complementary to perceived caregiver burden. The problem with self reports on actual workload – which is supposed to provide an objective measure – is that an element of subjectivity may influence them. For example, caregivers sometimes indicated they spent time caring for their relative 24 hours per day, 7 days per week, and at the same time spent 24 hours each day on household tasks, which is of course impossible. It is difficult to exclude this element of subjectivity from the self reports. Diaries may provide an alternative, but this requires much more commitment and compliance from already highly burdened people.

Economic analysis:

perspectives and assumptions

The perspective of the health care system instead of the wider and more time-elaborate societal perspective was judged to be acceptable, because the effect of our intervention on patients' and caregivers' productivity were judged to be negligible. We focused on all cost drivers relevant for the viewpoint taken, this meant that medication, laboratory and diagnostics costs were not taken into account, because the influence of our intervention on these cost parameters was judged to be negligible as well.

Comparison with other studies

Patient effects:

were our results clinically relevant?

The beneficial effects on disability and mental well-being represent a 5 and 10% better performance compared to control conditions, respectively. When comparing these results with the available literature we conclude that our results are of clinical relevance. Our results indicate that it is possible to prevent deterioration of functional skills for about three months and to improve well-being for at least half a year with a fairly simple home-based intervention for frail older patients. Nevertheless, the interpretation of study results in terms of clinical relevance is a very difficult issue if results are not overwhelmingly large.

The most recent meta-analysis of comprehensive geriatric assessment (CGA) showed benefits for Home Assessment Services on mortality, living at home, and on hospital admission, and for Outpatient Assessment Services on living at home. However, both showed no benefits for physical functioning [8]. The two recent meta-analyses on preventive home visits found beneficial yet non-significant, and not clinically relevant improvements in functional status [9, 10], and the standardised effect size [11] was around 0.05 [10]. The large MRC trial on population-based multidimensional assessment of older people in the UK found hardly any clinically relevant positive quality of life outcomes, but this may be related to the lack of a control group receiving no intervention in this trial [12].

The magnitude of our effects (approximately 5 to 10%) is in line with treatment effects on health-related quality of life of other positive studies incorporating a frail population. An evaluation of outpatient geriatric evaluation and management found favourable differences in SF-36 mental health and physical functioning scores of about five percent and two percent respectively at 12 months of follow-up [13]. A trial with a model of integrated care and case management for frail older people living in the community found favourable differences of 18.1% in basic activities of daily living, 6.9% in instrumental activities of daily living, and 6.8% in depression [14]. A home based intervention programme for physically frail older people found approximately 13% improvement in disability at seven months follow-up [15].

Further support for clinical relevance of our study results can be found if we recalculate our study results as standardised effect sizes (table 2). Then, we see that most outcome measures show beneficial effects above (and for the primary outcome even well above) the established threshold of 0.2 as the lowest threshold for a "minimally clinically important difference" [11]. Despite this criterion being arbitrary, it has been used in earlier interpretations: the MRC trial utilised the same threshold as a cut-off for clinical relevance [12, 16].

table 2 *Standardisations of the observed difference: Cohen standardised effect size[11] and the effect as a fraction of the score range*

	Observed difference (standardised effect size)*	Observed difference (fraction of score range)†
Groningen Activity Restriction Scale-3 3mths	0.40	0.06
Groningen Activity Restriction Scale-3 6mths	0.24	0.04
Medical Outcome Scale-20 subscale mental health 3mths	0.33	0.06
Medical Outcome Scale-20 subscale mental health 6mths	0.42	0.09
Dementia Quality of Life negative affect 3 mths	0.42	0.04
Dementia Quality of Life negative affect 6 mths	0.24	0.03
Dementia Quality of Life positive affect 3 mths	0.24	0.03
Cantril's self anchoring ladder 3 mths	0.21	0.04
De Jong-Gierveld Loneliness scale 3 mths	0.04	0.01
Medical Outcome Scale-20 subscale physical functioning 3 mths	0.21	0.04
Medical Outcome Scale-20 subscale role functioning 3 mths	0.11	0.05
Timed Up and Go Test 3 mths	0.10	0.03
Mini Mental State examination 6 mths	0.14	0.02

* Cohen standardised effect size is calculated as treatment difference/standard deviation

† The fraction of score range is calculated by dividing the measure through the appropriate score range, for the Timed Up and Go test it was divided through 60 seconds (because Timed Up and Go test has no score range)

The recent LOTIS trial also used GARS as the primary outcome measure [17]. They used GARS with four (GARS-4) instead of three (GARS-3) answer categories and only 13 out of 18 items [18, 19]. This means that the score can range between 13 and 52. A difference of 3 points was taken to be the detectable difference in the sample size calculations,

assuming a standard deviation of 7.5. In terms of score range this difference is 8%. Our minimally detectable difference was substantially higher: 12,5%. Some other studies applied GARS, though these did not report the minimally detectable differences used [20-22].

The comparison with other positive studies, the recalculations of our results as standardised effect sizes, and the comparison with other studies applying GARS, support our conviction that our results are of clinical relevance. However, as in most other studies, the results are not overwhelmingly large. In this situation, interpretation of the research results in terms of clinical relevance becomes a difficult and sometimes arbitrary matter. Frequent use of fewer instruments (to gain more experience with an instrument), input from workers in the field and considering the results in relation to patient and caregiver opinions is very important.

Caregiver effects

The effect of geriatric models aimed at the patient has not been consistently studied with regard to caregiver effects. The literature that is available suggests potential benefits for the caregivers [23-26]. Our results suggest that potential benefits are possible, but that the effect of intervening at the level of the patient on caregiver burden is not the same for all caregivers.

This observation probably relates to the heterogeneity among caregivers in our study. Caregivers who lived together with the patient were substantially more burdened, and seemed to benefit from this intervention, while caregivers living separately did not benefit. Our caregivers who shared a household were older, and more frequently men than the caregivers living separately. Other have suggested that the benefits of intervening relate to the gender of the caregiver, and that female caregivers sharing a household with the patient may benefit the most [27].

Our observations require further evaluation in larger studies, but lend support to the assumption that caregiver characteristics are important. Not every caregiver is the same, and these characteristics determine the feeling of being burdened as well as the possibilities to deal with this. As society's policies regarding population ageing increasing rely on informal care [28], this may have important public health consequences.

Selection procedure:

targeting the right patient

The evidence suggests that efficacy of geriatric intervention models including community-based models, probably benefits from targeting suitable patients [8, 9, 29, 30]. Targeting has therefore been advocated as a means to improve programmes' effectiveness by avoiding caring for patients unlikely to benefit [31, 32].

"Too ill" or "too healthy" are often heard characterisations of patients unlikely to benefit. The issue of appropriate targeting has more recently also been a matter of debate in relation to the introduction of Evercare in the UK [33]. Evercare identifies patients as suitable for intervention based on their past admissions record [34]. Population screening is another frequently used approach to targeting [35], but this is expensive and not easy to implement in daily practice. Generally speaking there is a broad spectrum of targeting approaches

from targeting on the basis of – predictive – personal characteristics measured with some sort of instrument (age, number and type of chronic diseases, sex, disability levels etc.) in a whole population to problem-based approaches to a large extent based on (implicit) clinical judgement. Different types of targeting often accompany different types of intervention. For example, a population screening approach is much more suitable for a programme of preventive home visits than for a geriatric evaluation and management unit. The targeting approach based on problems presented at a doctor and on clinical judgement is very close to everyday health practice.

Another predictor of adverse outcomes is the presence of geriatric syndromes [36]. To our knowledge, a geriatric intervention model with a problem-based type of patient selection has not been studied in a randomised way earlier. Our study adds an additional dimension to the debate on appropriate targeting by suggesting that specifically targeting frail older people with geriatric clinical syndromes can be an effective strategy.

Content of the intervention

The Dutch Geriatric Intervention Programme aimed to provide additional nursing assessment and nursing interventions. The intention of our model was that patients or caregivers clearly defined the main problem and goal of intervention in cooperation with their general practitioners in advance. The intervention would further focus on this aim. In practice, general practitioners often negotiated with their patients only a very loosely described aim. Therefore, the problem analysis also had a much wider perspective than only additional nursing assessment. As a consequence coordination of care also became a very prominent part of the interventions; fewer interventions focused, for instance, on psycho-education or therapy monitoring. More information on the content of our intervention model when used in practice is available in the thesis of van Eijken titled: "Strategies for improving community health care for the elderly [37]".

Future health care implications

Finding every patient:

how to accomplish this, and do we really want to?

There is another, more fundamental issue lying underneath the discussion on effective targeting strategies. Williamson et al. pointed at this already long ago when they noticed that "most old people do not report their complaints to their doctors until the condition is advanced" [38]. More than forty years and many health care system reforms later, this is still a major issue in geriatric care; why did we not recognise the problems of this patient at an earlier, more modifiable stage? Patient compliance is crucial to the success of screening based models to identify and handle unrecognised needs. Yet, it is often overlooked. Problem-based models have the commitment of the patient, but need to ensure that a considerable number of the potentially eligible patients are reached. Otherwise, there will be only individual, but not public health, benefits.

Often presented solutions to the problem of unmet needs are approaches based on population screening: identification of all persons using a systematic, more or less detailed screening method. Preventive home visits and health risk appraisal, and the annual health check by general practitioners in the UK are important examples [39-41]. The advantages of screening are obvious: no one is missed because of its systematic nature

and it provides timely intervention if done on a regular base. There are disadvantages as well: screening is costly, often inefficient, and may mean medicalisation of old age, although this last aspect is not necessary only negative [42]. The use of a previously defined set of standard criteria for diagnostic screening programmes [43, 44], reveals a crucial aspect that is often overlooked, and that is the aspect of patient adherence. Programmes can only be effective, if patients comply with the advice that results from them. In the more traditional fields of screening this may not be a problem: if a person is confronted with the diagnosis of cancer, this person is easily motivated to accept therapy. However, this can be different in our field of care; corrections of vision, hearing, or risks in the living environment do not carry the same absolute necessity as the treatment of a cancer. One can find several examples in literature, where patient non-compliance is suggested as an explanation for a negative study result [17, 45]. In another example – a randomised study of preventive screening for disorders among older people – an intervention was effectively started in less than half of the newly diagnosed disorders [46]. As a consequence, in order to succeed, time and effort has to be spent on seeking the commitment of the persons targeted. Many preventive health care programmes too easily take patient adherence for granted.

The problem of non-compliance may be dealt with, if we are able to identify persons who have that commitment. This can be done, for example, by intervening on conditions they (or their doctors) experience as a problem. This characteristic of intervention – solicited versus unsolicited – might be an important determinant of successful intervention. Our study is an example of the solicited approach. Despite the warning comments of Williamson et al. [38], we have shown that problem-based solicited intervention does not come too late. It may in fact be better timed, and has the commitment of both the general practitioner and the patient, and may therefore be more effective.

However, some caution is needed when interpreting our results in terms of potential benefits for public health. The intervention showed benefits for the individual, but can it also improve the health of the population of older people as a whole? Important in this discussion are the generally low numbers of patients general practitioners recruited for this study. Especially if the general practitioner overlooks a lot of potentially treatable persons, this would be a fundamental barrier for being of help to all – not just a few – older people. Our study does not provide the exact figures, but general practitioners seemed to be selective. We do not know whether this was intentional and based on valid arguments or because problems were unknown to them. Maybe, the results of a recent study on screening for disorders of older people by general practitioners are an indication that the latter is not such a problem: less than half of the disorders found on screening were regarded as new information by the general practitioners [46]. The intervention may also have been withheld from a patient for the wrong reasons, e.g. because of too pessimistic views on possible benefits. However, the general practitioners were able to select patients eligible for the intervention, not many of the patients who were discussed with the geriatricians for referral were rejected as inappropriate.

What next steps can be taken? We should make sure that screening based methods are reserved for disorders that satisfy the conditions identified for the conduct of an effective preventive screening programme, including the assurance of a sufficient level of compliance. In problem-based approaches we must ensure that all potential

beneficiaries are offered the opportunity to be treated. Here, there is also an important role for the responsible primary care professional. They need to have a sensitive and proactive attitude towards older patients' health disorders.

Integration versus fragmentation

If one wants to put the scientific knowledge from this study into practice, how should this be accomplished? Which type of nurses should participate and where should they be stationed? One of the criticisms which might be levelled against our model is that it leads to further fragmentation, whereas integration is highly needed. This model offers integration if general practitioners and nurses share responsibility for coordination of care for vulnerable older people. Unfortunately, in many models piloted throughout the Netherlands the role of general practitioners is less prominent than desirable.

In this project a specialist geriatric nurse based at a hospital conducted the home visits. If this model were to be introduced into primary care in exactly the same way, the criticism of creating more fragmentation might be justified. However, a study has other requirements than regular practice, such as the need for full accessibility of the participating professionals for data acquisition and uniform reporting. This was one of the reasons for assigning hospital nurses to do the home visits in this research project. Of course, implementation of this type of intervention in usual primary care needs translation and tailoring to that practical situation. Knowledge translation projects probably are an important vehicle to do this [47]. Knowledge translation "is defined as the exchange, synthesis and ethically sound application of knowledge – within a complex system of interactions among researchers and users – to accelerate the capture of the benefits of research through improved health, more effective services and products, and a strengthened health care system." The EASYcare/GIDS project is an example of how the knowledge generated by the Dutch EASYcare Study can be implemented in dementia care in general practice [48].

Although hospital nurses conducted the home visits in this study, primary care nurses (home care nurses, or practice nurses) might be the ideal professionals to carry out this intervention. In that situation, specialist geriatric nurses and geriatricians are experts who can provide teaching of primary care professionals and are available for consultation, a model which is very close to the so-called Quattromodel[49].

Preferably, we would entrust home care nurses (based in a locally organised home care service) and general practitioners (based in the general practice or primary care trust serving the same community) with the prime responsibility for geriatric primary care. Both are in a natural position to coordinate care of people with many co-morbidities [50, 51], and their activities are very much complementary in the care for older people. This would mean true integration and multidisciplinary care. However, it will require some fundamental choices in the organisation of health care to achieve this in the Netherlands. In the Netherlands, one-on-one cooperation between the general practitioner and home care nurse was lost in the increase in scale of home care organisations [52]. Though this probably happened unintentionally, we must to admit that this was a mistake and that one-on-one cooperation needs to be re-established.

Also, we should admit that – although the disease-centred approach of our health care system has been and still is very successful in fairly fit, independent, assertive, (cognitively) healthy adults experiencing one single problem – it can be a disaster for the vulnerable, cognitively impaired (older) adult with co-morbidity [53]. Vulnerable older people require a more holistic approach.

A necessary cautionary comment should be added about the general practitioner. An inventory by Van der Linden showed that home care organisations, hospitals and nursing homes have a more prominent role in transmural care projects than general practitioners [54]. Experiences with the Landelijk Dementie Programma (LDP; National Dementia Programme for the Netherlands for the improvement of dementia care in the Netherlands) [55] are divergent, but generally confirm this picture. Overall, the participation of general practitioners in the projects from LDP is smaller than hoped for. We think that in the Netherlands general practitioners' position in care networks is much less prominent than expected on the basis of their central position in primary care. Moreover, their voice in the debate on geriatric primary care is missed. A systematic analysis of the causes of the relative absence of the general practitioner (increasing work loads, time restraints, budgeting cuts, lack of organisation, lack of expertise, reluctance, wrong beliefs about possibilities, etc.) is needed, but beyond the scope of this thesis.

The current solution for these shortcomings in our health care system is to fix them with rather ad-hoc and uncoordinated initiatives like the introduction of practice nurses specialised in geriatrics or outreaching activities by nursing home physicians. However, a structural solution requires us to make more radical choices. (Re)establishing the close collaboration of home care nurse and general practitioner should be at the core of this solution. Some promising steps along this path have already been taken [56]. General practice is challenged to seriously invest in geriatric primary care, and make this a prominent issue, otherwise their position in an important and increasing area of health care is marginalised at the expense of further fragmentation of care.

Implications for future research

Future research should focus on the following subjects:

- Translation of these study results into clinical practice. How do we implement this knowledge?
- Replication of these study results in larger studies to determine which programme characteristics are effective. As the models implemented in clinical practice have to satisfy local conditions, it will be difficult to introduce uniform models on wide scale. If we know which elements are effective, we can build tailored models using these blocks.
- Further investigation is required into the role of different caregiver characteristics on caregiver burden, on the effect of intervention at a patient level on caregiver burden, as well as their influence on the outcome of interventions to decrease the burden of caring.

References

1. Ferrucci L, Guralnik JM, Studenski S, Fried LP, Cutler GB, Jr., Walston JD. Designing randomized, controlled trials aimed at preventing or delaying functional decline and disability in frail, older persons: a consensus report. *J Am Geriatr Soc* 2004 April;52(4):625-34.
2. van Iersel MB, Verbeek AL, Bloem BR, Munneke M, Esselink RA, Rikkert MG. Frail elderly patients with dementia go too fast. *J Neurol Neurosurg Psychiatry* 2006 July;77(7):874-6.
3. Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. *Stat Med* 2001 December 30;20(24):3931-46.
4. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res* 1999 March;8(1):3-15.
5. Heijmans MJWM, Rijken PM, Schellevis FG, van den Bos GAM, . Meer dan een ziekte. De gevolgen van comorbiditeit vanuit het perspectief van chronisch zieken en gehandicapten (More than just one disease. The consequences of comorbidity from the perspective of persons with a chronic disease or handicap). Utrecht: NIVEL; 2003.
6. Rockwood K, Howlett S, Stadnyk K, Carver D, Powell C, Stolee P. Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. *J Clin Epidemiol* 2003 August;56(8):736-43.
7. Rockwood K, Stadnyk K, Carver D et al. A clinimetric evaluation of specialized geriatric care for rural dwelling, frail older people. *J Am Geriatr Soc* 2000 September;48(9):1080-5.
8. Stuck AE, Siu AL, Wieland GD, Adams J, Rubenstein LZ. Comprehensive geriatric assessment: a meta-analysis of controlled trials. *Lancet* 1993 October 23;342(8878):1032-6.
9. Stuck AE, Egger M, Hammer A, Minder CE, Beck JC. Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis. *JAMA* 2002 February 27;287(8):1022-8.
10. Elkan R, Kendrick D, Dewey M et al. Effectiveness of home based support for older people: systematic review and meta-analysis. *BMJ* 2001 September 29;323(7315):719-25.
11. Cohen J. A Power Primer. *Psychological Bulletin* 1992 July;112(1):155-9.
12. Fletcher AE, Price GM, Ng ES et al. Population-based multidimensional assessment of older people in UK general practice: a cluster-randomised factorial trial. *Lancet* 2004 November 6;364(9446):1667-77.
13. Cohen HJ, Feussner JR, Weinberger M et al. A controlled trial of inpatient and outpatient geriatric evaluation and management. *N Engl J Med* 2002 March 21;346(12):905-12.
14. Bernabei R, Landi F, Gambassi G et al. Randomised trial of impact of model of integrated care and case management for older people living in the community. *BMJ* 1998 May 2;316(7141):1348-51.
15. Gill TM, Baker DI, Gottschalk M, Peduzzi PN, Allore H, Byers A. A program to prevent functional decline in physically frail, elderly persons who live at home. *N Engl J Med* 2002 October 3;347(14):1068-74.
16. Fletcher AE, Jones DA, Bulpitt CJ, Tulloch AJ. The MRC trial of assessment and management of older people in the community: objectives, design and interventions [ISRCTN23494848]. *BMC Health Serv Res* 2002 October 25;2(1):21.
17. de Craen AJ, Gussekloo J, Blauw GJ, Willems CG, Westendorp RG. Randomised Controlled Trial of Unsolicited Occupational Therapy in Community-Dwelling Elderly People: The LOTIS Trial. *PLoS Clin Trials* 2006 May;1(1):e2.
18. 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 December;43(11):1601-10.
19. Kempen GI, Doeglas D, Suurmeijer TP. *Het meten van problemen met zelfredzaamheid op verzorgend en huishoudelijk gebied met de Groningen Activiteiten Restrictie Schaal (GARS) (Groningen Activity Restriction Scale), een handleiding*. 2002.
20. Stiggelbout M, Popkema DY, Hopman-Rock M, de GM, van MW. Once a week is not enough: effects of a widely implemented group based exercise programme for older adults; a randomised controlled trial. *J Epidemiol Community Health* 2004 February;58(2):83-8.
21. Tak E, Staats P, Van Hespen A, Hopman-Rock M. The effects of an exercise program for older adults with osteoarthritis of the hip. *J Rheumatol* 2005 June;32(6):1106-13.
22. Faber MJ, Bosscher RJ, Chin A Paw MJ, van Wieringen PC. Effects of exercise programs on falls and mobility in frail and pre-frail older adults: A multicenter randomized controlled trial. *Arch Phys Med Rehabil* 2006 July;87(7):885-96.
23. Hung LC, Liu CC, Hung HC, Kuo HW. Effects of a nursing intervention program on disabled patients and their caregivers. *Arch Gerontol Geriatr* 2003 May;36(3):259-72.
24. Silverman M, Musa D, Martin DC, Lave JR, Adams J, Ricci EM. Evaluation of outpatient geriatric assessment: a randomized multi-site trial. *J Am Geriatr Soc* 1995 July;43(7):733-40.
25. Tibaldi V, Aimonino N, Ponzetto M et al. A randomized controlled trial of a home hospital intervention for frail elderly demented patients: behavioral disturbances and caregiver's stress. *Arch Gerontol Geriatr Suppl* 2004;(9):431-6.
26. Weuve JL, Boult C, Morishita L. The effects of outpatient geriatric evaluation and management on caregiver burden. *Gerontologist* 2000 August;40(4):429-36.
27. Vernooij-Dassen M, Huygen F, Felling A, Persoon J. Home care for dementia patients. *J Am Geriatr Soc* 1995 April;43(4):456-7.
28. WHO. Active ageing: a policy framework. Geneva: World Health Organisation; 2002.
29. Rubenstein LZ, Stuck AE. Preventive home visits for older people: defining criteria for success. *Age Ageing* 2001 March;30(2):107-9.
30. Wieland D. The effectiveness and costs of comprehensive geriatric evaluation and management. *Crit Rev Oncol Hematol* 2003 November;48(2):227-37.
31. Rubenstein LZ, Goodwin M, Hadley E et al. Working group recommendations: targeting criteria for geriatric

evaluation and management research. *J Am Geriatr Soc* 1991 September;39(9 Pt 2):37S-41S.
32. Winograd CH. Targeting strategies: an overview of criteria and outcomes. *J Am Geriatr Soc* 1991 September;39(9 Pt 2):25S-35S.
33. Roland M, Dusheiko M, Gravelle H, Parker S. Follow up of people aged 65 and over with a history of emergency admissions: analysis of routine admission data. *BMJ* 2005 February 5;330(7486):289-92.
34. Implementing the Evercare programme. Interim report. London: United Health Group; 2004.
35. Wieland D, Rubenstein LZ. What do we know about patient targeting in geriatric evaluation and management (GEM) programs? *Aging (Milano)* 1996 October;8(5):297-310.
36. Winograd CH, Gerety MB, Chung M, Goldstein MK, Dominguez F, Vallone R. Screening for Frailty - Criteria and Predictors of Outcomes. *Journal of the American Geriatrics Society* 1991 August;39(8):778-84.
37. van Eijken MI. Strategies for improving community health care for the elderly. Radboud Universiteit Nijmegen Medical Centre; 2007.
38. Williamson J, Fisher M, Stokoe IH et al. Old People at Home - Their Unreported Needs. *Lancet* 1964;283(7343):1117-20.
39. Stuck AE, Beck JC, Egger M. Preventing disability in elderly people. *Lancet* 2004 November 6;364(9446):1641-2.
40. Stuck AE, Elkuch P, Dapp U, Anders J, Iliffe S, Swift CG. Feasibility and yield of a self-administered questionnaire for health risk appraisal in older people in three European countries. *Age Ageing* 2002 November;31(6):463-7.
41. Department of Health United Kingdom. Terms of service for doctors in general practice. London: HM Stationary Office; 1989.
42. Ebrahim S. The medicalisation of old age. *BMJ* 2002 April 13;324(7342):861-3.
43. Bouter LM, van Dongen MCJM. *Epidemiologisch onderzoek: opzet en interpretatie (epidemiological research: design and interpretation)*. 3, herzien ed. Houten/Diegem: Bohn Stafleu Van Loghum; 1995.
44. Wilson JMG, Jungner G. Principles and Practice of Screening for Disease. *Who Chronicle* 1968;22(11):473-&.
45. Gussekloo J, de Bont LEA, von Faber M et al. Auditory rehabilitation of older people from the general population - the Leiden 85-plus Study. *British Journal of General Practice* 2003 July;53(492):536-40.
46. Eekhof J, De Bock G, Schaapveld K, Springer M. Effects of screening for disorders among the elderly: an intervention study in general practice. *Fam Pract* 2000 August;17(4):329-33.
47. Davis D, Evans M, Jadad A et al. The case for knowledge translation: shortening the journey from evidence to effect. *BMJ* 2003 July 5;327(7405):33-5.
48. Perry M, Draskovic I, Lucassen P et al.; The design of the GIDS-study: A randomised controlled trial on the effectiveness of a dementia training programme for general practitioners and practice nurses on dementia diagnosis, management and collaboration in primary care. 2007.
49. Beusmans G, Crebolder HF, van Ree J. Zorg voor chronisch zieken: Praktijkverpleegkundige breed inzetten. *Medisch Contact* 2001 February 16;56(7):259-62.
50. Schellevis FG. Je gaat het pas zien als je het door hebt: multimorbiditeit in de huisartsenpraktijk (oratie). 2006. Utrecht, NIVEL.
51. *Primary care in the driver's seat? Organizational reform in European primary care*. Maidenhead, Berkshire: Open University Press; 2006.
52. van Oorschot JMP. *Kruiswerk cruciaal : geschiedenis van de Kruisvereniging Zuid-Gelderland 1926-1996*. Nijmegen: Kruisvereniging Zuid-Gelderland; 1996.
53. Tinetti ME, Fried T. The end of the disease era. *Am J Med* 2004 February 1;116(3):179-85.
54. van de Linden BA. The birth of Integration: explorative studies on the development and implementation of transmural care in the Netherlands 1994-2000. University of Utrecht; 2001.
55. Meerveld J, Schumacher J, Krijger E, Bal R, Nies H. Landelijk Dementieprogramma. Werkboek. Utrecht: Nederlands Instituut voor Zorg en Welzijn / NIZW; 2006.
56. Landelijke Eerstelijns Samenwerkings Afspraak Decubitus - M82. *Huisarts wetenschap* 2004;47(13):652-4.

Summary

There is a need for effective geriatric primary care because vulnerable older people will increasingly live independently. Unfortunately, the studies on the effectiveness of geriatric primary care models such as geriatric evaluation and management, preventive home visits, and intermediate care have given conflicting results. The Dutch Geriatric Intervention Programme is a nurse-led, multidisciplinary programme of house calls for vulnerable older people. This study evaluates aspects of feasibility, effectiveness and cost-effectiveness of this type of health care for vulnerable older people.

Chapter 1 describes the context against which this thesis is set. The chapter describes how vulnerable older people are in need of special attention and why there is the need for a major reform of health care for older adults. From literature we conclude that the effectiveness of different intervention models for vulnerable older people has not been convincingly shown.

Moreover, effective methods are needed to study the effects of different types of health care for vulnerable older people. Certain methodological problems are important in the context of our study: problems related to studying older people such as high attrition rates, study subjects heterogeneity, and problems related to the selection of the best outcome measures. Other problems are related to studying a primary care population and studying complex interventions. A crucial problem related to studying complex interventions is that very often their evaluation cannot be done blind. As a result contamination bias and selection bias may occur. In our study, both contamination and selection bias threatened to harm its valid conduct when the regularly available randomisation methods for individual and cluster randomisation were used. Pseudo cluster randomisation was developed to deal with this dilemma.

Chapter 2 applies the example of intermediate care to discuss disagreeing definitions and how this hampers scientific appraisal of the available knowledge. As there is little scientific evidence on the benefits of intermediate care, research is especially important. A prerequisite for research is agreement on the definition of a concept, which is lacking for intermediate care. Commonly used definitions of intermediate care do not help much, because several very different definitions are in use. What is needed at the outset is a consensus on what constitutes intermediate care. The aim of this debate should not be to arrive at a uniform definition of intermediate care, for our inventory on the definitions of intermediate care shows that it is impossible to define intermediate care unequivocally at the highest conceptual level. For reasons of simplicity, this debate should be limited to defining intermediate care for the purpose of scientific appraisal. For the time being we believe that intermediate care models can be best classified according to their objectives of care and not by their names.

Chapter 3 is a detailed description of the design of our main study. The background section identifies elements from the different models available in literature that may be related to increased effectiveness. Such elements are: a multidisciplinary approach, longer and more intensive follow up, clinical control over the implementation of recommendations, and involvement of primary care providers. It further points at the debate on participation selection: should populations be selected or unselected, older or younger, healthy or vulnerable? In the methods section, the principles of the Dutch Geriatric Intervention Programme are described. The Dutch Geriatric Intervention Programme used a problem-based selection performed by the general practitioner rather

than population screening to identify patients eligible for participation. The problems targeted concerned cognition, nutrition, behaviour, mood, or mobility, and had to require nursing assessment, coordination of care, monitoring of therapy, or case management. Up to six visits by a specialist geriatric nurse for additional geriatric evaluation and management were planned within a three months period. The nurse conducted the main part of the intervention. The general practitioners continued their usual medical care and intervened in individual cases as agreed upon during interdisciplinary consultations with the nurse and geriatrician. The general practitioner kept prime responsibility for the care the patient and made the final decisions. We developed guidelines for each of the five presenting health problems to structure activities, without losing flexibility in tailoring individual interventions.

The **chapters 4 and 5** discuss pseudo cluster randomisation: chapter 4 describes the method, and chapter 5 the results of an evaluation of the performance when we used pseudo cluster randomisation in the Dutch EASYcare Study. Pseudo cluster randomisation was used because individual randomisation risked contamination and cluster randomisation risked selection bias due to expected treatment preferences of participating general practitioners. Pseudo cluster randomisation is a two-stage randomisation procedure that first randomises general practitioners in two groups (H and L). Then, patients of one general practitioner are randomised in majority to the intervention (intervention:control = 80:20) or to the control arm (intervention:control = 20:80). Selection bias and enrolment problems in the control groups are prevented because patients are randomised individually. Contamination occurs less because exposure to the intervention treatment is concentrated in a group of physicians recruiting almost exclusively participants who will be included in the intervention group (like cluster randomisation). While the assumptions underlying pseudo cluster randomisation sound reasonable, they need empirical evaluation. Part of this evaluation could be done using data from the Dutch EASYcare Study. This evaluation showed that general practitioners had strong treatment preferences which would have affected recruitment if cluster randomisation had been applied. There were some indications for the presence of contamination in case individual randomisation had been used: 65% of general practitioners who recruited more than one patient for our study and had patients in the intervention group stated that they had used elements of the intervention in the treatment of control participants.

After the recruitment period ended, a large majority of general practitioners estimated that equal numbers were allocated to both treatment arms, and if general practitioners estimated asymmetric randomisation ratios they tended to be more uncertain. We conclude that for cluster sizes of up to ten, asymmetrical randomisation in a 20:80 (= ratio intervention:control) or 80:20 did not lead to predictability of the randomisation sequence. The baseline characteristics of the study arms were highly comparable in the Dutch EASYcare Study. This gives no indication of selection bias. The variation in the number of patients each physician included explained the unbalanced numbers of control and intervention group. Overall, there were no indications for enrolment rate differences between study arms. We conclude that – when properly used – pseudo cluster randomisation provides a good alternative if individual randomisation risks contamination and cluster randomisation risks differential recruitment.

The **chapters 6, 7, and 8** present the main findings of the Dutch EASYcare Study. The effects on the patient's health-related quality of life, caregiver burden, and cost-effectiveness are discussed in the different chapters. The Dutch EASYcare Study followed 151 patients (mean age 82 years, 75% women) and 110 caregivers (mean age 56 years, 73% women, 58% offspring caregiver) for six months. 85 patients (and 61 caregivers) were randomised to the intervention programme, and 66 patients (and 49 caregivers) received usual care. Participants were very vulnerable: they had a mean Mini Mental State Examination score of 22 and a co-morbidity score (CIRS-G) of 10.1; the latter score is comparable to that of patients admitted to the geriatric ward of the Radboud University Nijmegen Medical Centre. After three months, patient outcomes showed significant differences in favour of the new intervention. Functional abilities (GARS-3) improved 2.2 points (95% confidence interval: 0.3 to 4.2) and well-being (MOS-20 mental health) 5.8 points (0.1 to 11.4). After six months the favourable effect increased for well-being (9.1 (2.4 to 15.9)), but the effect on functional abilities was no longer significant (1.6 (-0.7 to 3.9)). DGIP survival at two years was higher (82% versus 73%), although statistical significance was not reached (Log-Rank test p=0.40). Overall, perceived caregiver burden showed no significant differences over time; at three months this difference was -0.67 (-4.0 to 2.7), and 2.29 (-1.6 to 6.2) at six months. However, perceived burden was at baseline more than eight points higher in caregivers sharing a household with patients (n = 23) compared to caregivers living separately (n = 87). The intervention performed convincingly better in caregivers living together with the patient than in caregivers living separately (p for interaction = 0.04). Possibly, this difference is explained by baseline differences in perceived burden of both subgroups: the intervention decreased caregiver burden in the most burdened group of caregivers.

The average cost of the intervention under study was 998 euros [95% CI 888 – 1108]. The increment in total costs resulting from the new intervention was little over 761 euros [-3336 – 4687]. Hospitalisation and institutionalisation cost less; home care, adult day care, and meals-on-wheels costs were higher. To obtain an interpretable effect measure, we considered the treatment a success if a patient's MOS-20MH score increased more than 10 points and the GARS-3 score declined no more than 4.5 points. We used this responder definition because the course of GARS-3 scores showed a decline in our control group, while fairly steady MOS-20MH scores were observed over a six months period. Relative to the score range, both cut-offs represent approximately 10% change in score. There was a significant difference in the proportions of successful treatments of 22.3% [4.3 – 41.4]. The number-needed-to-treat was approximately 5 [2.3 – 18.0]. The incremental cost-effectiveness ratio is 3418 euros per successful treatment [-21458 – 45362]. The new treatment is cost-effective at a willingness-to-pay of 34000 euros.

The final **chapter 9** summarises the results and puts them in a broader methodological, scientific, and societal context. We conclude that problem-based nursing care effectively supported primary health care for vulnerable older people, and did so at reasonable cost.

Despite the careful design of this study, several methodological aspects deserve consideration. The most important are the validity of the application of pseudo cluster randomisation – an issue that has already been addressed in this summary – and the effect of attrition. To study the effect of attrition on the outcomes, we performed

a sensitivity analysis assigning the "mean of the other group" to the missing values, which is a conservative method. The results of the sensitivity analyses showed the same picture as the primary analyses. A drawback is that imputing the mean of the competing group seems rather irrational from a clinical point of view. However, the use of other strategies like "last-observation-carried-forward" or multiple imputation have stronger disadvantages.

An important issue concerns the clinical relevance of these results. The introduction of this thesis already pointed at the difficulties involved in finding valid outcome measures that were applicable to a very diverse, yet very frail, population of older adults, and were also well-known. This also complicates the interpretation of study results in terms of clinical relevance, especially if the results are not overwhelmingly large. There were beneficial effects on patients' disability (5% better performance) and mental well-being (10% better performance). When comparing these results with the available literature we conclude that our results are in the same range as benefits that were judged to be of clinical relevance previously.

Popular targeting strategies which have been used in the past are population screening and selection, or targeting patients who have been hospitalised previously. Our study adds an additional dimension to the debate on appropriate targeting by suggesting that specifically targeting frail older people with incident geriatric clinical syndromes can be an effective strategy. However, it is important to remember that problem-based models need to ensure that a considerable number of the potentially eligible patients are reached. Otherwise, there will be only individual, but not public health benefits.

If one wants to put the scientific content of this study into practice, how should this be accomplished? Which type of nurses should participate in this model and where should they be stationed? One of the criticisms which might be levelled against our model is that it leads to further fragmentation, whereas integration is highly needed. This model offers integration if general practitioners and nurses share responsibility for coordination of care for vulnerable older people. Unfortunately, in many models piloted throughout the Netherlands, the role of general practitioners is less prominent than desirable. This needs improvement. Everyone should be convinced of the importance of the general practitioner – as the gatekeeper to our health care system – in the care of community-dwelling older people, certainly when these older people are vulnerable.

Samenvatting

Kwetsbare ouderen zullen in de toekomst meer nog dan nu al het geval is zelfstandig willen leven. Trouwens, de maatschappij verwacht het eigenlijk ook van ze, want actieve deelname aan het maatschappelijke leven door zoveel mogelijk ouderen wordt gepresenteerd als een belangrijke oplossing voor het probleem van de vergrijzing. Om dit te bereiken is er veel behoefte aan goede en zo nodig gespecialiseerde zorg voor kwetsbare ouderen dicht bij huis. Er wordt dan ook door veel mensen gezocht naar de beste manieren om kwetsbare ouderen die zorg te bieden, want dat de huidige (ziektegerichte) zorg niet zondermeer aan de vereisten voldoet is al lang geen verrassing meer. Helaas is het juiste recept voor zorg voor kwetsbare ouderen die zelfstandig wonen nog niet gevonden. Er wordt nagedacht over bijvoorbeeld preventieve huisbezoeken of consultatiebureaus voor ouderen, maar de resultaten van het onderzoek naar de waarde ervan leveren een tegenstrijdig beeld op. Het zogenaamde Dutch Geriatric Intervention Programme (Nederlands Geriatrisch InterventieProgramma) is een model waarin huisartsen een gespecialiseerde verpleegkundige kunnen inschakelen in de zorg voor kwetsbare ouderen in hun praktijk. Dit proefschrift, waarin de resultaten van de Dutch EASYcare Study worden beschreven, laat zien dat deze nieuwe vorm van zorg voor kwetsbare ouderen ertoe bijdraagt dat het welzijn van de ouderen verbetert en hun zelfredzaamheid behouden blijft. Bovendien blijkt de nieuwe zorg kosteneffectief ("het product biedt waar voor zijn geld").

Om de meerwaarde van de verpleegkundige huisbezoeken betrouwbaar te kunnen beoordelen, zijn in de Dutch EASYcare Study twee groepen met elkaar vergeleken. De helft van de deelnemende ouderen werd wel door de verpleegkundige bezocht en de andere helft niet. Die beide groepen moeten natuurlijk goed vergelijkbaar zijn om een eerlijke vergelijking te kunnen maken. Daarvoor worden de deelnemers door randomisatie (loting) verdeeld over de groepen. Met de bestaande randomisatiemethoden was het echter niet goed mogelijk om tot een eerlijke verdeling van de deelnemers over de groepen te komen. Daarom ontwikkelden we een nieuwe methode om proefpersonen aan onderzoeksgroepen toe te wijzen. Deze nieuwe methode heet pseudoclusterrandomisatie. Wanneer en hoe deze nieuwe randomisatiemethode kan worden gebruikt en of de methode ook voldoet aan de verwachtingen is het tweede belangrijke onderdeel van dit proefschrift.

Appels met peren vergelijken is vaak een hachelijke zaak. Immers, appels en peren zijn niet vergelijkbaar. Evenzo is het in sommige gevallen onwenselijk dat appels en peren onder het mom "het is allemaal fruit" op een grote hoop geschoven worden. Hetzelfde geldt in wetenschappelijk onderzoek. Vergelijken is belangrijk, maar dan moet wel duidelijk zijn waar we precies over spreken. Goede definities van de zorgvormen die aan bod komen is dus van groot belang. Aan de hand van het voorbeeld van de zogenaamde Intermediate care modellen – dit zijn nieuwe zorgvormen die tussen de eerstelijns zorg en de ziekenhuiszorg in staan – laat **hoofdstuk 2** zien dat het vaak ontbreekt aan heldere definities over wat we precies onder een bepaalde zorgvorm verstaan. Er bestaan heel veel verschillende definities van het begrip Intermediate care. Dit alles maakt de discussie over de meerwaarde van dit veelbelovende type zorg erg moeizaam. Om toch een zinnige wetenschappelijke discussie te kunnen voeren stellen we voor om zorgvormen vooral te vergelijken aan de hand van het doel dat ze voor ogen hebben en minder aan de hand van hun naam.

Een belangrijk thema dat in **hoofdstuk 3** aangesneden wordt is de vraag hoe te voorkomen dat de groeiende groep kwetsbare ouderen door de mazen van het zorgstelsel heen valt, zoals dat nu soms wel gebeurt. De grofweg meest voorgestelde oplossing is om

heel systematisch alle ouderen op de aanwezigheid van kwetsbaarheid en nog niet behandelde problemen en ziekten te screenen. Voor gevonden problemen wordt vervolgens een oplossing gezocht. Hoewel dit heel aantrekkelijk lijkt, wordt duidelijk dat deze aanpak nog veel haken en ogen kent: moeten echt alle ouderen gescreend worden of kan beter een selectie gemaakt worden? Welke selectie moet dat dan zijn? Moeten dat nog relatief gezonde ouderen zijn, of moeten het juist meer kwetsbare ouderen zijn? Verder beschrijft hoofdstuk 3 de opzet van de studie en het model van verpleegkundige huisbezoeken dat onderzocht werd. Een belangrijk verschil tussen het door ons gebruikte model van verpleegkundige huisbezoeken en andere modellen is dat wij ervoor gekozen hebben om de huisarts de huisbezoeken te laten opstarten. De huisarts startte de huisbezoeken op het moment dat sprake was van belangrijke problemen zoals geheugenproblemen, sombere stemming of vallen. Het grote voordeel van deze aanpak is dat de verpleegkundige komt op het moment dat er iets aan de hand is. Andere belangrijke voordelen zijn dat de huisarts, de patiënt en de familie gemotiveerd zijn om voor de problemen een oplossing te zoeken. Een nadeel is dat mensen die niet met hun problemen bij de huisarts komen op deze manier niet geholpen kunnen worden.

De methode van pseudoclusterrandomisatie wordt besproken in de **hoofdstukken 4 en 5**. In hoofdstuk 4 wordt de methode uitgelegd en in hoofdstuk 5 wordt beschreven of de methode naar behoren werkt. Voorzover te beoordelen blijkt dat het geval te zijn. De pseudoclusterrandomisatie methode kan gebruikt worden wanneer het gebruik van clusterrandomisatie tot onvergelijkbare groepen leidt en het gebruik van individuele randomisatie juist tot contaminatie. Contaminatie zou in de situatie van randomisatie van individuele patiënten (individuele randomisatie) waarschijnlijk optreden, omdat de huisarts met de verpleegkundige samenwerkte in de uitvoering van de huisbezoeken. Het is dan niet uit te sluiten dat de huisarts elementen van de huisbezoeken ook zou gaan gebruiken bij deelnemers waarbij de huisarts juist op de oude voet door moet gaan. Onvergelijkbaarheid zou juist ontstaan als we huisartsen met al hun deelnemende patiënten zouden randomiseren (clusterrandomisatie), hetgeen de gangbare oplossing is om contaminatie te beperken. Dit zou kunnen optreden, omdat de huisarts wist of zijn patiënten wel of niet door de verpleegkundige bezocht zouden worden. Die kennis kon de keuze van deelnemers door de huisarts verschillend beïnvloeden, omdat niet alle deelnemers op hetzelfde moment geïncludeerd werden.

De belangrijkste resultaten van de Dutch EASYcare Study staan in de **hoofdstukken 6, 7 en 8**. Er zijn 151 deelnemers die – naar verwachting – erg kwetsbaar waren zes maanden gevolgd: 85 deelnemers werden bezocht door een verpleegkundige en bij 66 deelnemers ging de huisarts op de oude voet verder.

Na drie maanden was het behandeleffect voor zelfredzaamheid 2,2 punten [95% BI 0,3–4,2] en voor welbevinden 5,8 punten [0,1–11,4] ten voordele van de nieuwe interventie. Na zes maanden bleven de verschillen bestaan (zelfredzaamheid 1,6 [-0,7–3,9]; welbevinden 9,1 [2,4–15,9]), alleen voor zelfredzaamheid niet langer significant. Gemiddeld veranderde de mantelzorgerbelasting van de 110 (niet bij iedere deelnemer was een mantelzorger aanwezig) deelnemende mantelzorgers niet significant. Er was echter een significante interactie tussen behandeling en woonsituatie van deelnemer en mantelzorger (p=0,04): mantelzorgers die met de patiënt samenwoonden leken door de huisbezoeken minder belasting te ervaren dan mantelzorgers in de controlegroep die

samenwoonden, terwijl mantelzorgers die apart woonden mogelijk een verslechtering van de belasting ervoeren ten gevolge van de huisbezoeken.

De verpleegkundige huisbezoeken kostten ongeveer 1000 euro per deelnemer. De totale zorgkosten gedurende zes maanden waren ongeveer 750 euro hoger: zoals gehoopt werden minder kosten gemaakt voor ziekenhuisopname of opname in een verzorgings- of een verpleeghuis en kostten thuiszorg, deeltijdbehandeling en tafeltje-dek-je meer. De verpleegkundige huisbezoeken leverden 22 procent [4,3 – 41,4] meer "succesvol" behandelde deelnemers op, de number-needed-to-treat was ongeveer 5 [2,3 – 18,0]. "Succesvol" behandeld waren die deelnemers bij wie de zelfredzaamheid niet achteruitging, terwijl het welbevinden verbeterde. De incrementele kosteneffectiviteitsratio bedroeg 3418 euro per extra succesvol behandelde deelnemer [-21458 – 45362].

Hoofdstuk 9 begint met de vaststelling dat dit type van verpleegkundige huisbezoeken een verbetering van de zorg voor kwetsbare thuiswonende ouderen kan betekenen en plaatst de resultaten in een breder kader. Het hoofdstuk gaat in op de wetenschappelijke beperkingen en maakt de vergelijking met eerder uitgevoerd wetenschappelijk onderzoek.

De verpleegkundige huisbezoeken hadden een positief effect op zelfredzaamheid en psychisch welbevinden van kwetsbare ouderen dat ook aanwezig blijft als de verpleegkundige de zorg weer helemaal teruggeeft aan de huisarts. Dit lijken bescheiden effecten, maar ze zijn wetenschappelijk en klinisch relevant: het betreft een groep zeer kwetsbare ouderen bij wie bijkomende gebeurtenissen zoals een longontsteking eenvoudig een sterke en vaak blijvende achteruitgang in de kwaliteit van leven opleveren. GARS-3 meet 18 (I)ADL-items ([Instrumentele] Activiteiten van het Dagelijks Leven) op een 3-punts schaal: 1. zonder moeite zelfstandig, 2. met moeite, maar zelfstandig en 3. niet zelfstandig. Twee punten verbetering betekent dus op twee van de 18 items een verbetering van "niet zelfstandig, score 3" naar "met moeite, maar zelfstandig, score 2" of van "met moeite, maar zelfstandig, score 2" naar "zonder moeite zelfstandig, score 1", of op één item van "niet zelfstandig, score 3" naar "zonder moeite zelfstandig, score 1". In de praktijk betekent dit dat het mogelijk is de natuurlijke achteruitgang in zelfredzaamheid van deze kwetsbare ouderen voor tenminste drie maanden te voorkomen, zonder dat deze verbetering na stoppen van de huisbezoeken direct verloren gaat. Tegelijkertijd neemt het welbevinden ook op zes maanden vervolgens nog duidelijk toe.

De studie maakt duidelijk dat het effect van de interventie op mantelzorgerbelasting afhangt van de leefsituatie van patiënt en mantelzorger. Het effect van het interventieprogramma op de mantelzorgerbelasting is aanmerkelijk beter voor mantelzorgers die samenwonen met de patiënt. Deze groep ervaart overigens ook duidelijk meer belasting dan de mantelzorgers die gescheiden wonen. Het mogelijk negatieve effect op mantelzorgers die niet samenwonen met de patiënt wordt wellicht verklaard door het feit dat deze interventie juist een verhoogd beroep op de hen doet.

Een andere verklaring kan zijn dat zij zich – doordat zij daarmee door de interventie geconfronteerd werden – meer bewust werden van de kwetsbaarheid en beperkingen van hun naaste. Gezien de toenemende nadruk die ook in het Nederlandse zorgstelsel gelegd wordt op mantelzorg, is dit een resultaat dat verder onderzocht moeten worden.

Een belangrijke vraag is uiteraard hoe de resultaten van dit onderzoek in de praktijk gebracht kunnen worden. We denken bijvoorbeeld aan uitvoering van de verpleegkundige huisbezoeken door de wijk- of de praktijkverpleegkundige in samenwerking met de huisarts. Op die manier kom je ook tegemoet aan het mogelijke commentaar dat dit alleen maar tot verdere fragmentatie van de zorg voor kwetsbare ouderen leidt. Daarvoor moeten huisartsen zich wel actiever en zichtbaarder met de zorg voor kwetsbare ouderen gaan bemoeien, in de praktijk van alledag, maar ook bij het nadenken over de inrichting van de zorg. Het belang van de huisarts in de zorg voor ouderen is groot, eerst en vooral als het om kwetsbare ouderen gaat.

Dankwoord

Dank jullie wel

Jaren heb ik naar dit moment toe gewerkt en dan is het voor mijn gevoel toch nog plotseling zover dat ik mijn werk, ons werk voor iedereen mag verdedigen en aan iedereen mag laten zien waaraan ik met zovelen gewerkt heb. Het is dan ook eerst en vooral een gevoel van trots waarmee ik naar de verdediging toeleef. Tegelijkertijd is er ook nu weer een gevoel van onzekerheid en de mensen die mij kennen zullen begrijpen wat ik bedoel. Dit zal ongetwijfeld nog de nodige zenuwen opleveren, maar ook daaraan zal ik het hoofd wel weten te bieden. Niet in de laatste plaats juist met dank aan jullie.

Zoals ieder project kent ook dit project een "eigenaar", zoals dat naar goed marketinggebruik heet. Ik noem dat met opzet, want ik ken weinig onderzoekers die zo goed aanvoelen wat de maatschappij van onderzoekers verwacht en dat weten te vertalen in goed wetenschappelijk onderzoek. Ik ben mijn eerste promotor, professor Marcel Olde Rikkert, erg dankbaar dat hij mij vroeg en dat hij mij steeds de ruimte heeft gegeven om het onderzoek op mijn manier uit te voeren, ook als hij mijn werkwijze misschien wat al te grondig vond. Beste Marcel, dank je wel voor je vertrouwen in mij en voor de mogelijkheden die je mij biedt om me steeds verder te ontwikkelen. Toen ik besloot mijn koers te wijzigen en mij vooral te richten op een wetenschappelijke carrière in plaats van onderzoek en patiëntenzorg te combineren, was dat misschien niet wat jij voor mij voor ogen had, maar in plaats van me te bedanken voor betoonde diensten en me voor het overige het allerbeste te wensen, zag jij de mogelijkheden van mijn betrokkenheid als onderzoeker binnen jouw afdeling.

Mijn tweede promotor is professor Theo van Achterberg en hij stond gedurende mijn promotie weliswaar wat meer op de achtergrond, maar hij speelde absoluut niet de tweede viool. Beste Theo, ik waardeer je vooral om je heldere analyses die de discussie over mijn werk regelmatig een cruciale wending in de goede richting gaven. Dank je wel daarvoor. Bij de opzet van de studie zag ik een groot methodologisch dilemma en niet alleen hielp mijn co-promotor dr. George Borm mij daar betrokkenen die het probleem niet zo zagen van overtuigen, hij voorzag het probleem ook van een oplossing. En dat, George, kenmerkt je voor mij het beste: als iemand die problemen van onderzoekers op liefst zo simpele wijze van een oplossing voorziet. Keep it simple! George, dank. De multidisciplinaire aanpak die in de verpleegkundige huisbezoeken centraal stond weerspiegelde zich ook in de multidisciplinaire samenstelling van de groep die deze huisbezoeken geëvalueerd heeft. Daarin was dr. Eloy van de Lisdonk het huisartsgeneeskundig geweten. Beste Eloy, dank je wel voor inbreng. Ik ben blij dat ik jou als co-promotor vanuit de huisartsgeneeskunde naast me heb staan.

Mijn sparring partner vanuit de Verpleegwetenschappen in dit project was dr. Monique van Eijken. Samen met Monique heb ik dit project mogen uitbouwen tot de succesvolle studie die zij – denk ik – geworden is. Dat hebben we met veel plezier gedaan. Monique, wij verschillen op sommige punten ontzettend van elkaar, maar niet in onze liefde voor de oudere mens. Getty Huisman werd gevraagd om even bij te springen om Monique's zwangerschapsverlof op te vangen. Het werd duidelijk heel wat meer dan dat. Getty, dank je wel. Andere collega's van de verplegingswetenschap, dank je wel voor jullie belangstelling voor en meeleven met het onderzoekswerk van Monique en mijzelf.

Met niemand heb ik zo intensief over methoden kunnen overleggen als met dr. Steven Teerenstra en volgens mij kan niemand zo precies wiskunde vertalen naar de dagelijkse (onderzoeks)praktijk. Beste Steven, dank je voor al je deskundige geduld met een statistische leek.

En dan het onderzoeksobject zelf… want dat waren Hanny Hordijk en Marleen Lenkens: de interventie bestond feitelijk uit twee verpleegkundigen waarvan de werkzaamheid wetenschappelijk getoetst werd. Maar zij waren natuurlijk veel meer. Beste Hanny en Marleen, het was voor jullie even wennen, zo'n rol in wetenschappelijk onderzoek, waarbij varen op je verpleegkundige intuïtie en expertise alleen niet mocht, maar waarbij je je ook steeds moest afvragen of het de wetenschappelijke evaluatie niet in weg stond. Mede dankzij jullie expertise kon het project slagen. Het resultaat mag er wezen: jullie zijn twee van de weinige verpleegkundigen die echt kunnen zeggen dat zij wetenschappelijk bewezen effectief werken! Als dat geen "evidence based nursing" is. Dank jullie wel.

Many other contributors to the Dutch EASYcare Study deserve to be acknowledged: professor Stuart Parker who helped me so much understanding the complexity of geriatric health care, and dr. Anders Wimo who contributed to the economic paper. Verder wil ik noemen dr. Eddy Adang die me de beginselen van het gezondheidseconomisch onderzoek heeft bijgebracht en sinds wiens begeleiding ik weet wat een "stochast" is. Dr. Michel Wensing, professor Myrra Vernooij, Hans Bor, Henk van den Hoogen en Hans Wolters voor de inbreng van hun expertise en constructieve kritiek. Niet te vergeten zijn de studenten die me als wetenschappelijk stagiaire of studentassistent geholpen hebben: Mebeline Boon, Dagmar Oude Lansink, Dagmar Klaassen, Mirjam Louws en Sonja Vliek. Op deze plek wil ik als laatste professor Willibrord Hoefnagels nog noemen; hij stak me aan met het geriatrische virus en daar ben ik nog altijd blij om.

De "onderzoekerskamer" – waar ik veel van mijn tijd gesleten heb – heeft op vele plekken gezeten en even zovele bewoners gekend. Daarmee wisselde de atmosfeer, maar nooit de collegialiteit of de stress van een deadline. Toen ik de kamer deelde met Jaap Remmen, Lilian Vloet, Willy Colier en Anja Bos was de beroering soms groot; ging het niet om het indienen van een artikel, dan ging het wel over opvoedperikelen of politiek. Ook latere kamergenoten trokken in en weer uit: Anke Persoon, Lia Middeljans-Tijssen en Jurgen Claassen en dat zijn ze nog vast niet allemaal. Met Arenda van Beek, Miriam Reelick, Miriam Faes en Marieke Perry is de sfeer tegenwoordig weer studentikoos en uitgelaten op de – ondertussen – twee onderzoekerskamers. De diversiteit van de "bewoners" blijft ondertussen groot: Leny Theunisse, Jan Oudenes, sinds kort ook Els Meeuwsen en nog een hoop min of meer tijdelijke verblijvers. Dank jullie allemaal voor jullie hulp, steun en gezelligheid. Eén collega onderzoeker verdient het om apart genoemd te worden en dat is dr. Marianne van Iersel, want veel van mijn ervaringen deel ik met haar. Ik vind het jammer dat het uitwisselen van de gedoetjes van alledag nu niet meer zo gemakkelijk is, als toen we nog samen op de "onderzoekskamer" zaten. Lieve Marianne, dank je wel.

De klinisch geriaters, arts-assistenten, verpleegkundigen en alle andere afdelingsmedewerkers dank ik voor hun belangstelling en meedenken. Het belang van de betrokkenheid van jullie bij het wetenschappelijk onderzoek is groot: als uitvoerders van onderzoek, als leveranciers van deelnemers, als afnemers van de resultaten, maar vooral ook voor het genereren van vraagstellingen en onderzoeksideeën. Alleen dan gebeurt er onderzoek waarop de "kliniek" zit te wachten.

In onderzoek moeten vaak vele praktische zaken gedaan worden, daarbij kon ik steeds weer rekenen op de hulp van Gemma, Cynthia en de andere secretariaatsmedewerkers. Voor de laatste loodjes kon ik gelukkig ook rekenen op de hulp van Lian van Druenen die ondanks alle drukke bezigheden weer tijd (net als bij de geboortekaartjes van onze dochters, proefschrift Maureen etc.) heeft gevonden om een prachtig ontwerp voor mijn proefschrift te maken. Dank jullie wel.

Tijdens het begin van mijn onderzoeksperiode miste ik als min of meer enige epidemioloog op de afdeling Geriatrie soms een epidemiologische omgeving, dat werd echter al snel gecompenseerd door Gerhard Zielhuis, die mij begeleid heeft tijdens mijn opleiding Epidemiologie A, en in de uurtjes junior refereren, waarbij er heerlijk epidemiologisch gehaarkloofd kon worden.

Een wetenschappelijk onderzoek als dit kan enkel worden uitgevoerd als er deelnemers zijn. En deelnemers kreeg ik niet vanzelf, daar had ik huisartsen voor nodig. Vele huisartsen hebben met veel kritische betrokkenheid geprobeerd om aan het onderzoek bij te dragen: dank jullie wel. Al zijn er deelnemers, dan nog kan de studie alleen succesvol worden als de deelnemende patiënten en hun familieleden bereid zijn een hoop tijd te steken in het invullen van lange vragenlijsten. Toch deden ze dat steeds met veel inzet en betrokkenheid. Terwijl ik hen dus steeds hartelijk wilde bedanken voor hun medewerking, kreeg **ik** als dank niet alleen ingevulde vragenlijsten terug, maar ook prachtige verhalen, Merci, Indische hapjes en bovendien werd ook mijn grammatica door deelnemers – en terecht – hier en daar bijgespijkerd. Maar nu heb ik echt het laatste woord: dank **jullie** wel.

Wat zou een mens moeten als het leven alleen maar uit werken bestond. Ik prijs me gelukkig met de vele vrienden uit grofweg "Boekel", "Nijmegen", "rest van Nederland" en "België". Het zijn er teveel om op te noemen, maar ik ben steeds weer blij met jullie vriendschap. Vriendschap waarbij ik mezelf kan zijn in plaats van een uitgebreid CV met academische titels en publicaties. Het is fijn om je op deze manier steeds weer te mogen realiseren dat er ook nog iets anders dan p-waardes en Hb-gehaltes bestaat! Bregje Oostvogels – die altijd aan een half woord genoeg heeft om mijn werkperikelen te begrijpen – wil ik wel even apart noemen. Bregje, ik ben blij dat jij mij paranimf wilt zijn.

De meest vertrouwde omgeving is die van mijn familie. Altijd weer kan ik rekenen op steun die vanzelfsprekend gevonden wordt, zonder dat daar veel woorden voor nodig zijn. Nergens kan ik meer mezelf zijn dan in mijn eigen "nest". Frank, Wilma, Edwin, Hans, Mariska, Marcel, Wendy, Willemijn, Ton, Maria, Bregje, Kerstin, Patrick, Mark, Bram, oma Alwine en mijn neefjes en nichtjes: ik ben blij met jullie. Mijn vader en moeder verdienen een speciaal plekje, want zij stimuleren mij – steeds nog – het meest.

Ik ben blij dat jullie ervoor gezorgd hebben dat ik op de MAVO terecht kon, om van daar uit – onder jullie vleugels – verder te groeien naar waar ik nu sta. Betere raadgevers kan ik mij niet wensen.

En dan zijn daar sinds enkele jaren mijn lieve meisjes Veerle en Lize. Nergens zie ik zoveel levenslust en nergens ook ontleen ik zoveel levensvreugde aan. Veerle en Lize: als ik met jullie speel en knuffel, weet ik zeker dat ik dat ik de juiste keuze gemaakt heb om in mijn carrière niet alles te willen. Lieve Maureen, ruim een derde deel van mijn leven zijn wij nu bij elkaar en het leven heeft ons aan elkaar vast geklonken. En er is niets waarmee ik zo gelukkig ben dan dit…Wij. Kus.

Iedereen nogmaals… **dank jullie wel!**

List of publications

Publications

Melis RJF, Adang E, Teerenstra S et al. The cost-effectiveness of a multidisciplinary intervention model for community-dwelling frail older people. **J Gerontol A Biol Sci Med Sci 2007 in press**

Melis RJF, van Eijken MIJ, Teerenstra S et al.; A randomised study of a multidisciplinary program to intervene on geriatric syndromes in frail older people who live at home (Dutch EASYcare Study) [Clinicaltrials.gov Identifier NCT00105378]. **J Gerontol A Biol Sci Med Sci 2007 in press**

Melis RJF, Teerenstra S, Olde-Rikkert MG, Borm GF. Pseudo cluster randomization performed well when it was used in the Dutch EASYcare Study. **J Clin Epidemiol 2007 in press**

Teerenstra S, Moerbeek M, **Melis RJF**, Borm GF. A comparison of methods to analyse continuous data from pseudo cluster randomized trials. **Stat Med 2007 Sep 30; 26(22): 4100-15**

Van Eijken MIJ, **Melis RJF**, Wensing M, Olde Rikkert MGM, van Achterberg T. Feasibility of a new community based geriatric intervention programme: an exploration of experiences of GPs, nurses, geriatricians, patients, and caregivers. **Dis Rehab 2007 in press**

van Geel WJA, Abdo WF, **Melis RJF**, Williams S, Bloem BR, Verbeek MM; A more efficient enzyme-linked immunosorbent assay for measurement of α-synuclein in cerebrospinal fluid. **J Neurosci Methods 2007 in press**

Reynish E, Cortes F, Andrieu S, Cantet C, Olde Rikkert M, **Melis RJF**, Froelich L, Frisoni GB, Jonsson L, Visser PJ, Ousset PJ, Vellas B. The ICTUS Study: A Prospective Longitudinal Observational Study of 1,380 AD Patients in Europe. Study Design and Baseline Characteristics of the Cohort. **Neuroepidemiology 2007 Sep 24;29(1-2):29-38**

Teerenstra S, **Melis RJF**, Peer PG, Borm GF. Pseudo cluster randomization dealt with selection bias and contamination in clinical trials. **J Clin Epidemiol 2006 April;59(4):381-6**

van Poelgeest A, **Melis RJF**, Ritmeijer C, Olde Rikkert MG. De handelwijze bij acute levensbedreigende verstikking bij geriatrische en verpleeghuispatiënten [The handling of acute life-threatening choking in geriatric and nursing home patients]. **Tijdschr Gerontol Geriatr 2006 Sep;37(4):147-51**

Melis RJF, van Eijken MIJ, Borm GF, Wensing M, Adang E, van de Lisdonk EH et al. The design of the Dutch EASYcare study: a randomised controlled trial on the effectiveness of a problem-based community intervention model for frail elderly people [NCT00105378]. **BMC Health Serv Res 2005; 5(1):65**

Melis RJF, Olde Rikkert MG, Parker SG, van Eijken MIJ. What is intermediate care? **BMJ 2004; 329(7462):360-1**

Borm GF, **Melis RJF**, Teerenstra S, Peer PG. Pseudo cluster randomization: a treatment allocation method to minimize contamination and selection bias. **Stat Med 2005; 24(23):3535-47**

MM Rovers, **RJF Melis**, K Ingels, GJ van der Wilt, GA Zielhuis, P van den Broek. Aandoeningen en behandelingen op het gebied van de Keel-, Neus-, en Oorheelkunde bij kinderen in de eerste twee levensjaren [Complaints and therapies in the field of ENT medicine in children in their first two years of life]. **Ned Tijdschr KNO heelk 1999;5:107-14**