

University of Groningen

The association between implementation and outcome of a complex care program for frail elderly people

Ruikes, Franca G. H.; van Gaal, Betsie G. I.; Oudshoorn, Liset; Zuidema, Sytse U.; Akkermans, Reinier P.; Assendelft, Willem J. J.; Schers, Henk J.; Koopmans, Raymond T. C. M.

Published in:
Family practice

DOI:
[10.1093/fampra/cmz065](https://doi.org/10.1093/fampra/cmz065)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Ruikes, F. G. H., van Gaal, B. G. I., Oudshoorn, L., Zuidema, S. U., Akkermans, R. P., Assendelft, W. J. J., Schers, H. J., & Koopmans, R. T. C. M. (2018). The association between implementation and outcome of a complex care program for frail elderly people. *Family practice*, 35(1), 47-52.
<https://doi.org/10.1093/fampra/cmz065>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Health Service Research

The association between implementation and outcome of a complex care program for frail elderly people

Franca G H Ruikes^{a,*}, Betsie G I van Gaal^b, Liset Oudshoorn^a,
Sytse U Zuidema^c, Reinier P Akkermans^a, Willem J J Assendelft^a,
Henk J Schers^a and Raymond T C M Koopmans^{a,d}

^aDepartment of Primary and Community Care, Radboud University Medical Centre, Radboud Institute for Health Sciences, Nijmegen, The Netherlands, ^bRadboud University Medical Centre, Radboud Institute for Health Sciences, IQ Healthcare, Nijmegen, The Netherlands, ^cDepartment of General Practice and Elderly Care Medicine, University of Groningen, University Medical Centre, Groningen, The Netherlands and ^dJoachim and Anna, Centre for Specialized Geriatric Care, Nijmegen, The Netherlands.

*Correspondence to Franca G H Ruikes, Department of Primary and Community Care, Radboud University Medical Centre, Radboud Institute for Health Sciences, P.O. Box 9100, 6500 H.B. Nijmegen, The Netherlands; E-mail: franca.ruikes@radboudumc.nl.

Abstract

Background. Over the last 20 years, the effectiveness of complex care programs aiming to prevent adverse outcomes in frail elderly people has been disappointing. Recently, we found no effectiveness of the CareWell primary care program. It is largely unknown to what extent incomplete implementation of these complex interventions influences their outcomes.

Objective. To examine the association between the degree of implementation of the CareWell program and the prevention of functional decline in frail elderly people.

Methods. Quantitative process evaluation conducted alongside a cluster-controlled trial. Two hundred and four frail elderly participants from six general practitioner practices in the Netherlands received care according to the CareWell program, consisting of four key components: multidisciplinary team meetings, proactive care planning, case management and medication reviews. We measured time registrations of team meetings, case management and medication reviews and care plan data as stored in a digital information portal. These data were aggregated into a total implementation score (TIS) representing the program's overall implementation. We measured functional decline with the Katz-15 change score (follow-up score at 12 months minus the baseline score). The association between TIS and functional decline was analyzed with linear mixed model analyses.

Results. We found no statistically significant differences in functional decline between TIS groups ($F = 1.350$, $P = 0.245$). In the groups with the highest TISs, we found more functional decline.

Conclusion. A higher degree of implementation of the CareWell program did not lead to the prevention of functional decline in frail elderly people.

Key words: Activities of daily living, delivery of health care, frail elderly, health plan implementation, integrated, outcome and process assessment, primary health care.

Background

In the past 20 years, studies on complex care programs for frail elderly people have shown inconsistent and disappointing results on the prevention of functional decline (1,2). These have been appointed to the heterogeneity in care formats, professionals involved, outcome measures used and the setting and intensity of the interventions (1,2). Moreover, it is increasingly recognized that implementation fidelity, i.e. the degree to which the intervention was carried out as intended, can affect the intervention's outcomes (3).

Complex care programs comprise of multiple interacting components and require professionals and patients to change their behaviour (4). Moreover, they target several organizational levels, and necessitate flexibility and tailoring (4). These features cause complex programs to show great variation in their implementation (5). Therefore, it is important to interpret the outcome results of these programs in the light of their degree of implementation (6). Nowadays, process evaluations of complex interventions are common, especially in health promotion and public health domains (5). However, integrating implementation and outcome data in statistical analyses still is uncommon (6).

Recently, we published the negative results of the multicomponent CareWell primary care program, that aimed to prevent functional decline in community-dwelling frail elderly people (7). To interpret the lack of effectiveness, we performed this study to gain insight into the degree of implementation of the program. We hypothesized that a higher degree of implementation would be associated with less functional decline. The following research questions were addressed:

- To what extent was the CareWell primary care program implemented as intended?
- What is the association between the degree of implementation of the program and its primary outcome, i.e. (the prevention of) functional decline?

Methods

Study design and setting

In the Netherlands, general practitioners (GPs) provide continuous, person-centred care within a strong primary care setting. GPs often collaborate with practice nurses in the delivery of chronic care for the elderly (8). Moreover, elderly care physicians (ECPs), i.e. medical practitioners that have specialized as primary care experts in geriatric medicine, increasingly operate (as consultants) in primary geriatric care (9). However, coordination between GPs, other primary and specialist care providers, and home care and community services often is insufficient and fragmented (10).

Therefore, we developed the CareWell primary care program. It was implemented in six GP practices in Nijmegen, the Netherlands in a cluster-controlled trial of 12 months between September 2011 and September 2012; six control practices delivered usual care (7). The process evaluation was conducted alongside this trial.

The power calculation was based on the cluster controlled effectiveness trial: we calculated that we would be able to detect an effect size of >0.32 by including 50 participants in each cluster (total $n = 600$, assuming equal clusters), using a power of 80%, a two-sided alpha of 0.05, an assumed ICC of 0.01 and an expected loss to follow up of 35% (7).

Target population

All practices were instructed to include 50 frail participants ≥ 70 years within a limited 2-month inclusion period prior to the start of the

intervention period, with the use of the EASY-Care TOS instrument. First, GPs use prior knowledge to subdivide 'not frail' from '(possibly) frail' elders. The second step involves trained nurses to perform a comprehensive geriatric assessment of (possible) frail elders during a home visit. Then, GPs and nurses weigh all signs into a final frailty judgment (11). Exclusion criteria were institutionalization and/or critical or terminal illnesses. Details on recruitment were reported previously (7).

The intervention

The CareWell primary care program consisted of four key components: (i) multidisciplinary team (MDT) meetings, (ii) proactive care planning, (iii) case management and (iv) medication reviews.

Each practice assembled a MDT consisting of a GP(s), practice nurse(s) and/or community nurse(s), an ECP and a social worker with elderly care expertise. MDT meetings were supposed to be held every 4–8 weeks, and at least half-yearly per participant—more often if indicated. In addition, team members were able to communicate virtually through a secured web-based health and welfare information portal (12).

Tailor-made proactive care plans, based on the individual health-related problems and goals as assessed with the EASY-Care TOS, were formulated for each participant at the start of the intervention. A structured format including somatic, functional, psychological, social and communicative domains was used. Professionals were instructed to revise participants' care plans after discussion in a MDT meeting at least every 6 months, and to store the revised care plans, even when unchanged, in the information portal.

A case manager (nurse or social worker) was assigned to each participant. Case managers were responsible for coordinating, monitoring and evaluating proactive care planning and for the MDT planning. They were instructed to support participants' goal setting and self-management, and to actively maintain contact with participants (and informal caregivers) by telephone or home visits at least half-yearly.

Last, the GP and nurse were instructed to conduct a yearly medication review for each participant, in collaboration with a pharmacist.

All professionals attended two preparatory educational meetings and received written instructions, coaching on the job and help-desk support when needed. Professionals received financial reimbursement for time-investment and overhead costs.

Assessment of implementation fidelity

We developed a total implementation score (TIS) composed of the four components: (i) MDT meetings, (ii) proactive care planning, (iii) case management and (iv) medication reviews.

Data collection

All professionals were asked to fill in monthly time registration forms for individual patients. To stimulate uniformity in and compliance with time registrations, structured timesheets with written instructions were sent each month. Community nurses were already familiar with these time registrations, as they were required by their employer.

In scoring the delivery of MDT meetings and medication reviews, time registrations were used as a proxy, i.e. registered time for that component on a particular date was accounted for as 'delivery' on that date. In scoring proactive care planning, two investigators (FR and LO) independently assessed the care plan data as stored in the information portal. A care plan needed to contain a minimum of two health care problems with associated treatment goals and actions in order to count as a sufficient care plan. To be defined as a new

version of a care plan, additional problems needed to be included or pre-existing problems needed to be adjusted. Also, care plan revisions (independent of whether changes to the plan were made) 6 months after the last revision were counted as new care plans, assuming the revision was done in the half-yearly MDT. In scoring the delivery of case management, case managers were instructed to daily register their time spent per participant, in minutes.

Measurement of implementation fidelity of the key components

The TIS construction was based on consensus in the research group, consisting of experts in the field and a statistician, after extensive discussion prior to the availability of the study data and without an available theoretical framework: TIS was calculated by summing the implementation scores of the individual components, i.e. '1' indicating that the component was 'implemented as intended', '0' if not (Table 1).

For each participant, MDT meetings scored '1' when two or more meetings were held; '0' when less than two meetings were held. Proactive care planning scored '1' when two or more care plan versions were stored; '0' when less than two care plan versions were available. Case management activities were intended to be tailored to individual participants' needs; limits were thus not set beforehand. After finding a large spread in the overall minutes registered for case management activities, and acknowledging the importance of this component in the delivery of integrated care (13), we revised our theoretical construct and decided to add additional weight to this component. Case management activities were then scored as follows: '2' if median time or more was spent, '1' if less than median time was spent and '0' if no time was spent. According to the Dutch guideline 'Polypharmacy in the Elderly', a medication review is indicated for patients with polypharmacy, i.e. the use of five or more chronically prescribed drugs (14). Therefore, the first step in medication review was the identification of the participants with polypharmacy. For participants without polypharmacy, the medication review was then complete; these participants scored '1'. For participants with polypharmacy, a thorough review needed to follow, after which score '1' was appointed. Without this formal review, participants with polypharmacy scored '0'.

Measurement of the TIS

Total implementation score, reflecting the degree to which the intervention was implemented as intended, was calculated by summing the scores of the four components into a sum score ranging from zero to five; a higher score reflecting a higher degree of implementation (Table 1).

Baseline characteristics of the target population

Participant's baseline characteristics were measured at baseline and at follow-up after 12 months.

Data analysis

We calculated frequencies and means of participants' baseline characteristics, implementation of the key components and the TISs at practice and participant level. Between-practices differences in means were analyzed with ANOVA.

The association between participants' Katz-15 change scores (i.e. follow-up score minus baseline score) and TIS were analyzed with linear mixed model analyses. We performed a model with a random intercept, representing the clustering of participants in GP practices and all other variables fixed. Depending on the linearity of the relationship between the Katz-15 change scores and TISs, the TIS would be taken as a continuous or categorical variable in the model.

All statistical analyses were performed using SPSS version 20. Tests were considered significant at $P < .05$.

Results

Baseline characteristics of the target population

We included 287 participants in the intervention group and had a loss to follow-up of 83 out of the 287 participants in the program due to death (10.8%), institutionalization/hospitalization (9.1%) and unknown other reasons/lost to follow-up (9.1%) (7). This study included 204 (71.1%) participants, ranging from 28 to 41 participants per setting. Baseline characteristics are shown in Table 2.

Implementation fidelity of key components

MDT meetings

Overall, complete MDT meetings were organized at least twice for 47.5% of the participants, with a mean of 1.5 team meetings per participant (SD 1.2, range 0–6). The degree of implementation of MDT meetings in GP practices ranged from 24.4 to 67.9%, $P = 0.002$.

Proactive care planning

Of the 204 participants, 51.0% had at least two proactive care plans formulated. The mean number of care plans per participant was 1.7 (SD 1.3, range 0–6). The implementation degree of proactive care planning in GP practices ranged from 3.4 to 94.7%, $P < 0.001$.

Table 1. Data collection and measurement of implementation scores of key components and TIS.

Key component	Measure	Source	Score
Multidisciplinary team work	Frequency	Time registrations	<2 meetings = 0 ≥ 2 meetings = 1
Proactive care planning	Number of care plan versions	Information portal	<2 care plan versions = 0 ≥ 2 care plan versions = 1
Case management	Time invested (minutes)	Time registrations	No time = 0 < Median time = 1 ≥ Median time = 2
Medication reviews	Frequency	Time registrations	Polypharmacy-, review- = 0 Polypharmacy+, review- = 1 Polypharmacy-, review+ = 1 Polypharmacy+, review+ = 1
Total implementation score	—	—	= Sum of above scores

Table 2. Baseline characteristics of participants^a.

Characteristics	GP practice						P value	Overall (N = 204)
	1 (N = 29)	2 (N = 28)	3 (N = 38)	4 (N = 30)	5 (N = 38)	6 (N = 41)		
Age, mean (years)	81.8	80.8	81.4	83.7	82.7	83.6	0.17	82.4
Female sex, %	75.9	60.7	73.7	73.3	76.3	68.3	0.75	71.6
Living alone, %	62.1	46.4	86.8	76.7	57.9	68.3	0.01	67.2
Socioeconomic status score ^b , mean	0.3	1.4	1.5	0.8	-0.6	-0.6	<0.001	0.4
Low level of education, %	31.0	11.1	36.8	36.7	5.3	24.4	<0.001	24.1
Cognition score ^c , mean	5.6	7.5	5.6	9.4	5.0	4.9	0.018	6.2
Baseline Katz 15 score ^d , mean	5.3	6.0	4.5	5.2	4.6	4.1	0.053	4.9
EQ-5D+C ^e , mean	0.59	0.51	0.59	0.57	0.68	0.72	0.034	0.62
RAND-36 Mental health ^f	60.6	63.3	64.7	57.6	60.3	62.7	0.22	61.7
Presence of health-related limitations in social functioning ^g	42.9	53.8	52.6	82.8	62.2	75.6	0.010	124
Frailty index ^h , mean	0.39	0.43	0.37	0.40	0.36	0.32	0.041	0.37
Presence of care complexity	27.6	7.1	13.2	6.7	32.4	19.5	0.030	18.2

GP, general practitioner.

^aValues are expressed as numbers (percentage) unless otherwise indicated.

^bSocioeconomic status score was based on postal code areas (income, employment and education); higher score indicates more social disadvantage.

^cBased on a modified Mini-Mental State Examination (MMSE; range 0 to 28); higher score indicates more cognitive problems.

^dKatz 15 score (range 0–15); higher score indicates more dependence in (instrumental) activities of daily living.

^eEuroQol-5D+C (EQ-5D+C) scores (range -0.33 to 1.00); higher score indicates a higher health-related quality of life.

^fRAND-36 Mental Health (range from 0 to 100); higher score indicates better mental health.

^gBased on the social functioning subscale of the RAND-36. Answers dichotomized in 'absence of limitations' versus the other categories indicating 'presence of limitations'.

^hThe frailty index measures accumulation of deficits (scale 0–1); a higher index suggests a more frail status.

Case management

Overall, 153 participants (75.0%) received case management; at practice level, this ranged from 46.3 to 97.4%. A mean of 155.8 min (SD 264, range 0–1625 min) was spent per participant, with a median of 62.5 min. The mean time spent per participant differed between practices with a range of 66.6–310.4 min, $P < 0.001$.

Medication reviews

One hundred and forty nine (73.0%) of participants had polypharmacy; 116 (77.9%) of them received a thorough medication review. 147 (72.1%) participants scored one point, of which 92 (62.6%) had polypharmacy. The degree of implementation of medication reviews differed between practices with a range of 47.4–85.7%, $P = 0.001$.

Implementation of the complete program: TIS

The mean TIS at participant level was 3.0 (SD 1.2, range 0–5), with a between-practices range of 2.3–4.0, $P < 0.001$. This variation was mainly caused by differences in the implementation of proactive care planning and case management.

The implementation of the program's key components and the TISs are presented in [Table 3](#).

Association between TIS and primary outcome

No linear association between the TISs and the Katz-15 change scores was found; the difference between TIS groups was analysed with TIS included as a categorical variable. We found no significant difference in Katz-15 change scores between TIS groups ($F = 1.350$, $P = 0.245$), as shown in [Table 4](#). However, the effect sizes of the Katz-15 change scores in the groups with a TIS score of 3, 4 or 5 exceed the a priori calculated effect size of >0.32. Sensitivity analysis

with TISs dichotomised in low (0–1–2) and high (3–4–5) scores underlined these results (data not shown).

Discussion

To our best knowledge, this is the first study that developed a quantitative implementation score to measure the degree of implementation and study the association between implementation and outcome of a complex care program for frail elderly people. We found no statistically significant differences in functional decline between TIS groups. The degree of implementation differed significantly between practices, mainly due to variation in the implementation of proactive care planning and case management. In contrast to our hypothesis, a higher degree of implementation tended to be associated with an increase in functional decline.

Our results show that implementation of the (key components of the) CareWell program in everyday GP practices is feasible, but leaves room for improvement. The practice with the highest degree of implementation showed the (second) best implementation scores for all key components, with the exception of medication reviews. The practice with the lowest degree of implementation had an exceptionally low score for proactive care planning. Although most participants in this practice did have one or more care plan versions stored in the information portal, these were either not updated or did not meet the requirements to be counted as a sufficient care plan. On the contrary, the practice with the highest degree of implementation had an exceptionally high score for proactive care planning. The influence of time and organizational constraints might be substantial. Prior experience with the concept and assessment of frailty, as observed in the practice with the highest score for care planning, might facilitate implementation. Between practices, we found a large

Table 3. Delivery of key components, total implementation scores and primary outcome scores at practice and participant level^a.

	GP practice						Overall N = 204
	1N = 29	2N = 28	3N = 38	4N = 30	5N = 38	6N = 41	
Practice characteristics							
No. of professionals involved	7	5	9	8	10	6	
Key components							
Multidisciplinary team work ^a	51.7	67.9	65.8	56.7	55.3	24.4	0.002
Proactive care planning ^a	3.4	64.3	94.7	60.0	10.5	65.9	<0.001
CM score, %							<0.001
0 (no time)	17.2	14.3	2.6	23.3	31.6	53.4	25.0
1 (less than median)	55.2	3.6	7.9	23.3	47.4	14.6	25.0
2 (median or more)	27.6	82.1	89.5	53.3	21.1	31.7	50.0
CM minutes, mean	66.6	310.4	287.5	92.4	105.5	84.3	155.8
Medication reviews ^a	65.5	85.7	47.4	66.7	84.2	82.9	0.001
Complete intervention							
TIS, mean	2.3	3.5	4.0	3.1	2.4	2.5	<0.001
Primary outcome							
Katz 15 change score, mean ^b	0.55	0.79	0.92	0.83	0.66	1.15	0.83

GP, general practitioner; TIS, total implementation score.

^aValues are expressed as percentage 'delivered as intended' (i.e. a score of one point).

^bA higher Katz 15 change score indicates more functional dependence in (instrumental) activities of daily living.

^cP value of the difference in means between practices (ANOVA).

variation in minutes spent on case management. This is suggestive of intentional and purposeful tailoring to individual participants' needs (15). Although multidisciplinary guidelines for follow-up care were available, difficulties in their use, as well as time constraints might have hindered the implementation of case management activities (15). Moreover, individual professionals' skills and learning curves might have contributed to inconsistencies in the delivery and quality of case management activities, despite antecedent training and coaching on the job (15). The implementation of MDT meetings showed a large variation between practices, possibly due to time and organizational constraints. Moreover, lacking knowledge on each others' roles and expertise, as well as time needed to build trusting working relationships might hinder truly integrated team work (13). Our twelve month follow-up period might be too short to achieve this. On the other hand, it is possible that 'delayed delivery' in MDT meetings and care planning was interpreted as non-adherence, while these were in fact intentional, tailored deviations.

We need to consider some study limitations. First, the power calculation of this study was derived from the effectiveness trial (7). The absence of significant differences in Katz-15 change scores between TIS groups might therefore be due to a type-II error. Although not

Table 4. Association between total implementation score and primary outcome (Katz-15 change score).

TIS	No. of participants	Katz 15 change score ^a , estimated effect	SE	95% CI
0	4	0.50	0.92	-1.32 to 2.32
1	28	0.54	0.35	-0.15 to 1.23
2	41	0.46	0.29	-0.11 to 1.03
3	53	0.70	0.25	0.20 to 1.20
4	56	1.20	0.25	0.71 to 1.68
5	22	1.36	0.39	0.59 to 2.14

TIS, total implementation score.

^aA higher Katz-15 change score indicates more functional decline regarding (instrumental) activities of daily living.

statistically significant, the observed effects in the three highest TIS groups might have clinical relevance as they exceeded the a priori calculated effect size (16).

Second, our theoretical framework underlying the construction of the TIS was based on research team consensus after deliberate discussion prior to data analysis. No existing literature on the conceptualization of an implementation score of complex interventions was readily available. Although the validity of our construct cannot be validated into detail, we believe it has face validity. Third, time registrations were used as a proxy for the delivery of two of the four key components. Although community nurses are used to fill in time registrations as endorsements of their hours worked, the time registrations of the practice nurses and social workers might have been incomplete or inaccurate due to time constraints, as is known from literature (15,17). However, it is unlikely that this selectively influenced these professionals and caused bias. A fourth limitation is that we were not able to include qualitative implementation data, e.g. the quality of the delivery of the components, in the analysis (5). This would have further strengthened our findings.

In a recently published primary elderly care trial in the BMJ, implementation data were linked to outcome by using a dichotomy variable (18). We aimed to take these analyses a step further by constructing a more refined implementation variable. In contrast to our hypothesis, we found that a higher degree of implementation tended to be counter intuitively associated with increased functional decline. We speculate that the program led to an increased, timelier awareness of participants' health and care risks, resulting in an increase of purposefully tailored interventions directed at those participants that were most prone of functional decline. However, the fact that these tailored interventions did not prevent functional decline raises some concerns. First, we used the validated EASY-Care TOS to identify the frail participants. However, during the intervention period, professionals deliberately targeted their interventions to those participants at highest risk of functional dependence, i.e. confounding by severity. It is possible that the targeted participants were already too frail for the program to show measurable effects on daily functioning.

Conversely, the participants that were identified to be less prone of functional decline might have been more susceptible to respond to the program. Second, it is possible that more person- or goal-oriented outcomes, e.g. goal-attainment scaling, better capture the effectiveness of the heterogeneous and tailored interventions, that were aimed at a diversity of risk factors for functional dependence (19). Third, the follow-up period might have been too short for this complex program to be optimally implemented and thus achieve its optimal effectiveness. Our fourth concern refers to the evaluation of the degree of implementation of our complex CareWell program, with its four interacting adaptive components. As we standardized the minimum implementation requirements of the components, it is possible that the dynamics of our complex intervention were not fully captured (20). Moreover, the program was implemented in GP practices that are on their own turn complex settings, in which change in input often is disproportionally correlated to change in outcome. The validity of our TIS construct in the light of the complexity of our program and its setting remains unclear.

Conclusion

A higher degree of implementation of the CareWell program did not lead to the prevention of functional decline in frail elderly people.

Declarations

Funding: this trial was funded by the Dutch National Care for the Elderly Program, as initiated by the Netherlands Organization for Health Research and Development (ZonMW project no. 313050201). Health insurance companies (CZ and UVIT) and the municipality provided additional funding.

Ethical approval: the local ethics commission CMO region Arnhem-Nijmegen reviewed the study and concluded that formal ethical approval was not necessary.

Conflict of interest: none declared.

References

- Beswick AD, Rees K, Dieppe P *et al.* Complex interventions to improve physical function and maintain independent living in elderly people: a systematic review and meta-analysis. *Lancet* 2008; 371: 725–35.
- Eklund K, Wilhelmson K. Outcomes of coordinated and integrated interventions targeting frail elderly people: a systematic review of randomised controlled trials. *Health Soc Care Commun* 2009; 17: 447–58.
- Hasson H, Blomberg S, Dunér A. Fidelity and moderating factors in complex interventions: a case study of a continuum of care program for frail elderly people in health and social care. *Implement Sci* 2012; 7: 23.
- Craig P, Dieppe P, Macintyre S *et al.* Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008; 337: a1655.
- Carroll C, Patterson M, Wood S *et al.* A conceptual framework for implementation fidelity. *Implement Sci* 2007; 2: 40.
- Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. *BMJ (Clinical research ed)* 2006; 332: 413–6.
- Ruikes FG, Zuidema SU, Akkermans RP *et al.* Multicomponent program to reduce functional decline in frail elderly people: a cluster controlled trial. *J Am Board Fam Med* 2016; 29: 209–17.
- van Weel C, Schers H, Timmermans A. Health care in the Netherlands. *J Am Board Fam Med* 2012; 25 (suppl 1): S12–7.
- Koopmans RT, Lavrijsen JC, Hoek JF, Went PB, Schols JM. Dutch elderly care physician: a new generation of nursing home physician specialists. *J Am Geriatr Soc* 2010; 58: 1807–9.
- Boeckxstaens P, De Graaf P. Primary care and care for older persons: position paper of the European Forum for Primary Care. *Qual Prim Care* 2011; 19: 369–89.
- van Kempen JA, Schers HJ, Jacobs A *et al.* Development of an instrument for the identification of frail older people as a target population for integrated care. *Br J Gen Pract* 2013; 63: e225–31.
- Robben SH, Perry M, Huisjes M, *et al.* Implementation of an innovative web-based conference table for community-dwelling frail older people, their informal caregivers and professionals: a process evaluation. *BMC Health Serv Res* 2012; 12: 251.
- de Stampa M, Vedel I, Bergman H *et al.* Opening the black box of clinical collaboration in integrated care models for frail, elderly patients. *Gerontologist* 2013; 53: 313–25.
- Verduijn MM, Leendertse AJ, Moeselaar A, *et al.* Multidisciplinary guideline for polypharmacy in the elderly. *Huisarts Wet* 2013; 56(8): 414–9.
- Bleijenberg N, Ten Dam VH, Drubbel I *et al.* Treatment fidelity of an evidence-based nurse-led intervention in a Proactive Primary Care Program for older people. *Worldviews Evid Based Nurs* 2016; 13: 75–84.
- Walters SJ. Consultants' forum: should post hoc sample size calculations be done? *Pharm Stat* 2009; 8: 163–9.
- Muntinga ME, Van Leeuwen KM, Schellevis FG, Nijpels G, Jansen AP. From concept to content: assessing the implementation fidelity of a chronic care model for frail, older people who live at home. *BMC Health Serv Res* 2015; 15: 18.
- Metzelthin SF, van Rossum E, de Witte LP *et al.* Effectiveness of interdisciplinary primary care approach to reduce disability in community dwelling frail older people: cluster randomised controlled trial. *BMJ (Clinical research ed)* 2013; 347: f5264.
- Rockwood K, Howlett S, Stadnyk K *et al.* Responsiveness of goal attainment scaling in a randomized controlled trial of comprehensive geriatric assessment. *J Clin Epidemiol* 2003; 56: 736–43.
- Hawe P, Shiell A, Riley T. Complex interventions: how “out of control” can a randomised controlled trial be? *BMJ* 2004; 328: 1561–3.