International Psychogeriatrics (2015), 27:11, 1875–1892 © International Psychogeriatric Association 2015
doi:10.1017/S1041610215000927

Dementia care mapping: effects on residents' quality of life and challenging behavior in German nursing homes. A quasi-experimental trial

Martin Nikolaus Dichter,[1,2] Tina Quasdorf,[1,2] Christian Günter Georg Schwab,[1,2] Diana Trutschel,[1,2,3] Burkhard Haastert,[4] Christine Riesner,[1,2,5] Sabine Bartholomeyczik[2] and Margareta Halek[1,2]

[1] *German Center for Neurodegenerative Diseases (DZNE), Witten, Germany*
[2] *School of Nursing Science, Faculty of Health, Witten/Herdecke University, Witten, Germany*
[3] *Institute of Computer Science, Martin-Luther-University Halle-Wittenberg, Halle, Germany*
[4] *mediStatistica, Neuenrade, Germany*
[5] *Division "Nursing Science and Pedagogics" Ministry of Health, Equalities, Care and Ageing of the State of North Rhine-Westphalia, Düsseldorf, Germany*

ABSTRACT

Background: Person-centered care (PCC) is a widely recognized concept in dementia research and care. Dementia Care Mapping (DCM) is a method for implementing PCC. Prior studies have yielded heterogeneous results regarding the effectiveness of DCM for people with dementia (PwD). We aimed to investigate the effectiveness of DCM with regard to quality of life (QoL) and challenging behavior in PwD in nursing homes (NHs).

Methods: Leben-QD II is an 18-month, three-armed, pragmatic quasi-experimental trial. The sample of PwD was divided into three groups with three living units per group: (A) DCM applied since 2009, (B) DCM newly introduced during the study, and (C) a control intervention based on a regular and standardized QoL rating. The primary outcome was QoL measured with the Quality of Life-Alzheimer's Disease (QoL-AD) proxy, and the secondary outcomes were QoL (measured with QUALIDEM) and challenging behavior (measured with the Neuropsychiatric Inventory Nursing Home version, NPI-NH).

Results: There were no significant differences either between the DCM intervention groups and the control group or between the two DCM intervention groups regarding changes in the primary or secondary outcomes. At baseline, the estimated least square means of the QoL-AD proxy for groups A, B, and C were 32.54 (confidence interval, hereafter CI: 29.36–35.72), 33.62 (CI: 30.55–36.68), and 30.50 (CI: 27.47–33.52), respectively. The DCM groups A (31.32; CI: 28.15–34.48) and B (27.60; CI: 24.51–30.69) exhibited a reduction in QoL values, whereas group C exhibited an increase (32.54; CI: 29.44–35.64) after T2.

Conclusions: DCM exhibited no statistically significant effect in terms of QoL and challenging behavior of PwD in NHs. To increase the likelihood of a positive effect for PwD, it is necessary to ensure successful implementation of the intervention.

Key words: nursing homes, dementia, quality of life, person-centered care, dementia care mapping, challenging behavior, psychosocial intervention

Introduction

Maintenance and enhancement of QoL is the main objective of dementia care (Moyle *et al.*, 2007). Because dementia is not curable, non-pharmacological interventions, particularly psychosocial interventions, are widely recognized for achieving a satisfactory QoL for PwD. However, despite recommendations for the use of psychosocial interventions as an initial treatment (Bartholomeyczik *et al.*, 2007; Vernooij-Dassen *et al.*, 2010), little is known about their effectiveness in terms of the QoL of PwD. One promising approach is person-centered care (PCC), the goals of which are to achieve the best possible QoL for PwD and to reduce challenging behavior (Chenoweth *et al.*, 2009). The theoretical basis of PCC is Kitwood's social–psychological

Correspondence should be addressed to: Martin Nikolaus Dichter, German Center for Neurodegenerative Diseases (DZNE), Witten, Germany. Phone: +49 2302 926 253; Fax: +49 2302 926 239. Email: Martin.Dichter@dzne.de. Received 22 Dec 2014; revision requested 3 Feb 2015; revised version received 18 May 2015; accepted 25 May 2015. First published online 3 July 2015.

theory of personhood (Kitwood, 1997). This theory assumes that dementia-unfriendly environments – including negative staff attitudes, communication, and care practices – strongly influence how PwD experience discomfort. PCC encompasses valuing PwD and their caregivers, providing individualized treatment looking at the world from the perspective of PwD and creating a positive social environment. These concepts are bundled into the VIPS framework (V = valuing PwD, I = individualized care, P = understanding the world from the patient's perspective, and S = providing a social environment that supports the needs of the patients) (Brooker, 2004). Existing evidence regarding different interventions for the operationalization of PCC, e.g. staff teaching programs, organizational change programs based on the VIPS Framework, or the observational method DCM, has suggested improvement in resident and staff outcomes (Brooker, 2004; Cohen-Mansfield *et al.*, 2012). Moreover, PCC has become an important component of dementia guidelines (NICE, 2006; Aged Care Branch, 2011).

The DCM method is a structured, multicomponent intervention that aims to operationalize and advance PCC in institutional and day-care settings. From 2001 to 2005, DCM was updated to the current eighth version (Brooker and Surr, 2006). The results of three international trials regarding the effectiveness of DCM are currently available. A first cluster randomized controlled trial (cRCT) conducted in Australia compared a standardized PCC training session, one DCM cycle, and task-oriented care (Chenoweth *et al.*, 2009). This trial identified a significant reduction in the primary outcome, "agitation," in NH units that provided PCC or DCM compared with units that provided task-oriented care (usual care) after eight months. No effects for the secondary outcomes of "challenging behavior (overall)" and "QoL" were observed.

In contrast, two recent European studies identified no effect of DCM on their primary outcomes. In a Dutch cRCT (van de Ven *et al.*, (2013), DCM was compared with usual care in dementia special care units over a period of eight months. This study identified no effects on the primary outcome, "agitation," or the secondary outcome "QoL" of PwD but indicated a negative effect on the secondary outcome "challenging behavior (overall)."

A cRCT in Norwegian NHs yielded similar results (Rokstad *et al.*, 2013). In this study, the DCM method was compared with a PCC intervention based on the VIPS framework and with usual care supported by a DVD-based education program over 11 months (Rosvik *et al.*, 2011). No effect on the primary outcome, "agitation," was found. However, the authors observed that DCM had a positive effect on the secondary outcomes "challenging behavior (overall)" and "QoL" for PwD.

The reasons for these heterogeneous results are unclear. Unfortunately, no further information regarding the implementation adherence or factors that facilitated or hindered the implementations of the interventions in these three studies is available. Moreover, there is a lack of knowledge regarding the long-term effects of DCM usage over periods of longer than 11 months. Furthermore, the heterogeneous results highlight the importance of cultural adaptation and intervention testing in different cultures and healthcare systems. In Germany, the DCM method has been well known for several years. However, there have been no studies of the effectiveness of DCM for residents with dementia in NHs in Germany.

Therefore, the purpose of our pragmatic trial, Leben-QD II, was to study the effectiveness of the DCM method for residents with dementia and their care staff in German NHs, to determine possible long-term effects of DCM and to gather knowledge regarding facilitating factors and barriers for implementing DCM in a real-life setting (Halek *et al.*, 2013). We defined the QoL of residents as the primary outcome because nurses who deliver PCC attend to resident's holistic needs as they experience, e.g. transitions in health or environmental conditions (Meleis *et al.*, 2000). Based on this holistic scope, QoL is the major outcome of nursing, especially PCC (O'Rourke *et al.*, 2015). In addition, PCC supports understanding of need-driven challenging behavior. The secondary outcome "challenging behavior" enables comparison with the results of previous trials. In this paper, we report our results regarding the following research question: does the DCM method positively affect the QoL of PwD and reduce their challenging behavior?

Methods

Research design

This pragmatic quasi-experimental trial was based on a convenience sample of nine NHs (each with one nursing unit) and three measurement points T0, baseline; T1, after six months; and T2, after 18 months (Halek *et al.*, 2013). The care home provider was responsible for assignment of a nursing unit to one of three different groups. The interest and prior experience of each nursing unit in the respective group intervention were crucial for the group assignment. Group A consisted of three units experienced in the use of DCM since 2009. In group B, the DCM method was implemented for the first time. In group C, a regular standardized QoL rating

was integrated into the usual care provided. The group assignment occurred prior to the baseline measurement.

The Ethics Committee of the German Society of Nursing Science approved the study design and protocol in August 2010. For further information about the research design, see Halek *et al.* (2013).

Sample and inclusion criteria

The participating NH units had heterogeneous living arrangements and care concepts. Thus, depending on the NH structure, dementia special care units with up to 15 residents or integrative traditional units with up to 36 resident places were included. All of the NHs were managed by one care home provider. The target population was all residents with dementia from the participating units.

Inclusion criteria for residents with dementia were written informed consent provided by the legal representatives, a Functional Assessment Staging Test (FAST) (Reisberg, 2007) score ≥ 2 and being present over the last two weeks in the NH unit.

Evidence (van de Ven *et al.*, 2013) and our experience led us to predict a relatively high attrition rate for this type of population. To enable an intention-to-treat analysis after the long intervention period (18 months), we included PwD in an ongoing procedure. Thus, new residents of the participating units were included at T1 and T2 if they met the inclusion criteria.

Intervention

The implementation of the DCM method in groups A and B was coordinated and managed by one social worker qualified as DCM trainer who had been employed by the NH provider for several years. For group C, a nursing manager from one of the included NHs coordinated the implementation of the control intervention. In addition, one staff member from each participating NH was deployed for the management and coordination of the intervention. The starting point for all interventions, as described below, was directly after the baseline measurement.

Intervention group A

The three units that had prior experience with DCM, received two DCM cycles per year until 2009. The observations, data analysis, report writing, and feedback of the results were performed by external registered nurses or social workers who were qualified as DCM mappers with evaluator status. For the intervention period in this study, two interested members of each care team were trained as DCM basic users by the in-house DCM trainer during a three-day course (part of component 1). During the DCM basic user course, the participants were trained in, e.g. PCC and DCM coding principles and practice and providing feedback. The training was based on the official DCM manual developed at the University of Bradford translated into German language by DCM Germany. After the training, the DCM basic users enhanced their skills during 1.5 days of observation sessions supervised by the in-house DCM trainer. Our hypothesis was that a higher level of PCC would be observed in these three nursing units than in groups B or C because the care staff in group A was experienced in DCM. This group might provide additional information about possible long-term effects.

The six DCM components (Brooker and Surr, 2007; BSI - British Standards Institution and University of Bradford, 2010) were implemented as follows (see also Halek *et al.* 2013):

(1) Briefing and preparation of care staff and leadership based on a formal initial meeting for each care staff team at the beginning of the intervention, a newsletter with information about the progress of the project created by the in-house DCM trainer and written materials (e.g. folders and posters) that described the objective and content of the project and DCM.
(2) DCM observations: 5–8 hours of DCM observations performed simultaneously by two trained DCM basic users. The observations were based on a standardized coding of the well-being and behavior of PwD. Moreover, the interactions between staff and people with dementia were assessed.
(3) DCM data analysis and report writing by the DCM basic user: presentation of a written feedback report regarding the results of the observations within the first week after the DCM observations.
(4) Feedback of results to care staff and leadership during a formal care staff meeting. The feedback was based on the written report and performed by the DCM basic user. The in-house DCM trainer participated on the first two of the three feedback meetings for each nursing unit.
(5) Based on the DCM results and reflection of the care staff during the feedback session, a written action plan was created by the care staff of the nursing unit within eight weeks after the DCM observations.
(6) Realization of the action plan by the care staff of the nursing unit.

Intervention group B

Before our trial, the three units in group B were inexperienced in providing PCC based on the DCM method, but they expressed great interest in DCM. For the intervention period of this study, two members of each care team were trained as DCM basic users. The procedure of the DCM intervention was equivalent to the procedures used in group A.

Control group C

Because of the quality assurance criteria of the German medical advisory service of the statutory health insurance regarding the mandatory assessment of QoL in PwD during the care process (MDS, 2009) and the interest in PCC by all of the participating units, it was not possible to establish a usual care group as a control group. Therefore, group C received an intervention based on a regular and standardized QoL rating that was integrated into the usual care. The proxy measurement QUALIDEM (Ettema *et al.*, 2007; Dichter *et al.*, 2011) was used for the QoL rating. This intervention consisted of the following components:

(1) Briefing and preparation of care staff and leadership.
(2) A 1.5-hours training program for the care staff on the three units that included general information about QoL rating for PwD and about the use of QUALIDEM in particular.
(3) An initial QoL rating for all residents with dementia within three months after the training (October–December 2011). The QoL ratings were always based on the proxy rating agreement between two staff members.
(4) Subsequent QoL ratings were conducted if notable changes in the residents' needs were perceived. At least every six months, a new QoL rating was conducted. If necessary, the staff initiated case conferences based on the results of the QoL ratings.

Instruments

The data collection and DCM cycles were combined. Following the baseline measurement (T0) and the first DCM cycle, measurement T1 occurred after an intervention phase of six months. After T1, second and third DCM cycles were performed, with a subsequent follow-up measurement (T2) after a total intervention period of 18 months. All of the instruments were assessed at each of the three time points. The instruments were chosen based on their appropriateness for the target setting and population, their psychometric properties, and their feasibility. All of the instruments were completed by caregivers (registered nurses and nursing aids) who were members of the living unit's care staff. In addition, to facilitate collection of up-to-date information about the resident, the caregivers had to have been at work on most days within the two weeks prior to data collection to ensure a close relationship with the resident being assessed. The process was guided by external, trained study assistants to ensure standardized data collection. The study assistants received 5 hours of prior training regarding data collection provided by the research team. A comprehensive instruction manual regarding data collection and a telephone hotline were provided.

QoL was measured as the primary outcome using the proxy version of the QoL-AD scale. This scale consists of 13 items and results in a score that ranges from 13 to 52, with higher scores indicating a higher QoL (Logsdon *et al.*, 1999; Graske *et al.*, 2014b). As recommended by Logsdon *et al.* (2002), up to two missing values were replaced with the mean score of the remaining items. More than two missing values resulted in exclusion of the scale data at the respective measurement point. This procedure was used for all instruments.

The secondary outcomes were QoL and challenging behavior exhibited by residents with dementia. The residents' challenging behavior was assessed with the NPI-NH, which includes 12 domains (Wood *et al.*, 2000). For each of the 12 domains, the frequency, severity, and distress for the caregivers were assessed. Scores were calculated for each domain as frequency * severity (range 0–12), and a total score was generated by adding the individual domain scores (range 0–144). Higher scores indicated the presence of more challenging behavior.

QoL as a secondary outcome was assessed with QUALIDEM (Ettema *et al.*, 2007; Dichter *et al.*, 2013a; 2014), which allows measurement of up to nine dementia-specific QoL domains. QUALIDEM consists of one 37-item version for people with mild-to-severe dementia and a consecutive 18-item version for people with very severe dementia. The response options for all items are "never," "rarely," "sometimes," and "frequently," thereby resulting in an item score between 0 and 3. The instrument focuses on the psychosocial domains of QoL and is feasible for assessing QoL in NHs. The subscale sum scores were transformed to values between 0 and 100. In addition, a QUALIDEM total sum score was calculated and transformed to values between 0 and 100 (QUALIDEM(%) = QUALIDEM(sum score) * 100 / (3 * *n*(number of items of each subscale))). Higher scores indicate a higher QoL. According to the individual FAST score, the QUALIDEM version for people with mild-to-severe (37 items) or very severe dementia (18 items) was used. When the dementia severity deteriorated from mild-to-severe to very severe during the study, the QUALIDEM score was calculated based on the 18-item version.

Functional ability was assessed with the Physical Self-Maintenance Scale (PSMS), which results in a score between 0 and 30, with higher scores indicating a lower functional ability (Lawton and Brody, 1969). The FAST was used for assessment of dementia severity. This staging instrument yields scores that range from 1 to 7 (1: free of cognitive impairment; 2–6: very mild dementia to severe

dementia; and 7: very severe dementia) (Reisberg, 2007). Data regarding regularly prescribed psychotropic drugs and pain medications were collected from the residents' medication sheets using a self-developed electronic assessment tool (Witten Longitudinal Medication Collecting Tool, WILMER), which exhibited satisfactory feasibility during a pre-test conducted before the data collection. With WILMER, it is possible to collect data regarding prescribed drugs and aggregate these data based on the anatomical therapeutic chemical (ATC) classification provided by the World Health Organization (WHO, 2003) for each prescribed substance. Age, gender, dementia diagnosis, and care dependency level, as defined by the German long-term care insurance (0–3), were taken from the residents' care documentation.

Influences on the institutional level were assessed in relation to the organizational characteristics, e.g. staff/resident ratio and funding per resident, with the self-developed Dementia Institution Questionnaire (DIQ). The DIQ is divided into two parts. The first part consists of four items that evaluate the entire institutional structure and financial revenue. The second part targets the nursing unit and is divided into resident-specific factors (three items), structure (one item, "How many residents rooms are there in the nursing unit?"), and staffing, including employee turnover (five items, e.g. "What is the sum of advanced training days (8 hours) for all employees of the nursing unit?"). The response for all items is open ended. The questionnaire was completed by the coordinating staff member of the corresponding NH. The instrument allows a descriptive analysis to be performed separately for each item, and combinations of item data can be used to determine relationships, e.g. between the number of advanced training days and staff full-time equivalents. The dementia-friendliness of the milieu of the nursing unit was assessed with the self-developed Dementia Milieu Assessment (DMA). This is a standardized observation instrument that consists of 29 items. The items are divided into an environmental subscale (21 items, e.g. "Are there many different noises, such as constant background music, TV noise, and loud conversations?") and a psychosocial subscale (eight items, e.g. "Care staff interact (behavior and conversation) respectfully with each person with dementia"). Each item is measured using dichotomous response options (yes/no). The scores range from 0 to 21 (environmental domain) and from 0 to 8 (psychosocial domain). Higher scores indicate a more dementia-friendly environment. The DMA observations were conducted by an experienced registered nurse and dementia researcher over a 2-hours period of observation from 3 pm to 5 pm in the public space of each participating nursing unit. The single items of the DIQ and DMA were judged as feasible in a cross-sectional study called Leben-QD I (Dichter *et al.*, 2013a).

Data analysis

The characteristics of residents and homes were stratified by time and intervention group and were described by frequency tables and means ± standard deviations, depending on their distribution. Overall, the three treatment groups were compared using Fisher's exact test or analysis of variance (ANOVA) tests. These analyses were not adjusted for cluster correlations. For the primary and secondary outcomes, normal distributions were assumed. This assumption was checked graphically (not shown). To calculate the NPI-NH total score, a sensitivity analysis was performed after square root transformation (results not shown). The expected values of the outcomes, including separate 95% confidence intervals for the three observation times and the three treatment groups, were estimated and adjusted for cluster correlation and repeated measurements (per resident) by fitting linear-mixed models (Brown and Prescott, 2006). The outcomes were used as dependent variables. Random effects were clusters (= homes) and interactions between clusters and time (large variance component of interaction). Any repeated measurement was adjusted by covariance patterns (Brown and Prescott, 2006) with compound symmetry (CS) structure and degrees of freedom calculated using Satterthwaite's method. The fixed effects (independent variables) were time, treatment group, and interaction time* treatment group. Furthermore, an adjustment for time-dependent covariates was performed in additional models by including the PSMS score (continuous), FAST score (binary: 2–6 vs. 7), pain medication (yes vs. no), and NPI-NH total score (continuous) as independent variables in the model. The choice of these variables as covariates was based on the significant group difference for pain medication use at baseline and the results of a systematic review of the factors associated with the QoL of PwD (Beerens *et al.*, 2013). Overall, type-three tests were performed for the primary fixed effects time, treatment, and interaction time * treatment. The expected values of the outcomes at time-treatment strata were estimated using model-based least square means. For covariate adjustment, the overall means of the PSMS score, "FAST > 6" (prevalence), pain medication (prevalence), and NPI-NH total score were assumed for all strata. The model-based estimated time courses of the three main outcomes are presented graphically with 95% confidence intervals (adjusted for cluster correlation

Table 1. Baseline characteristics of residents with dementia, nursing staff, and dementia milieu based on the intention to treat

T0[a]	GROUP A (DCM EXPERIENCED)	GROUP B (DCM)	GROUP C (REGULAR QoL-RATING)	TEST RESULTS *p*-VALUE
Residents with dementia	$n = 41$	$n = 52$	$n = 61$	
Age, years	82.5(±6.8)	84.1 (±6.3)	82.6 (±9.2)	0.48[c]
Women	33 (80)	43 (83)	52 (85)	0.80[d]
Dementia diagnoses (yes)	39 (95)	50 (96)	53 (87)	0.18[c]
Functional assessment				
Staging				
2 to 6	23 (56)	32 (62)	43 (70)	0.30[d]
7	18 (44)	20 (38)	18 (30)	
Physical self-maintenance scale (6–30)	20.0 (±5.9)	20.4 (±5.1)	18.7 (±5.6)	0.22[c]
Care dependency levels[b]				
0	0 (0)	0 (0)	1 (2)	
1	9 (22)	11 (21)	23 (38)	0.13[d]
2	15 (37)	21 (40)	23 (38)	
3	17 (41)	20 (38)	13 (22)	
Use of psychotropic drugs				
Yes (one)	16 (39)	21 (40)	25 (41)	0.37[d]
Yes (more than one)	18 (44)	26 (50)	22 (36)	
No	7 (17)	5 (10)	14 (23)	
Use of pain medication				
Yes (one or more)	17 (41)	6 (12)	22 (36)	0.001[d]
Nursing staff characteristics	$n = 3$	$n = 3$	$n = 3$	
Full-time equivalent quantity/resident	0.6 (±0.0)	0.8 (±0.2)	0.6 (±0.1)	
Full-time equivalent quantity/days for advanced training (8 h) last quarter	6.3 (±1.9)	9.7 (±7.8)	7.7 (±2.8)	
Employees turnover – *n* (% of nursing staff *n*)[e]	5 (5)	3 (2)	7 (5)	
Dementia Milieu				
Environmental subscale (0–21)	15.0 (±2.6)	15.3 (±5.0)	12.0 (±4.4)	
Social psychological subscale (0–8)	4.3 (±3.1)	3.7 (± 4.0)	2.0 (±3.5)	

Data are reported as the mean (±SD) or number (%).
[a]Overall $n = 234$ residents (all measurement points).
[b]As determined by expert raters of the medical service of the statutory long-term care. One missing in Group C.
[c]ANOVA.
[d]Fisher's exact test.
[e]Nursing staff include registered nurses, social workers and nursing assistants (nursing assistants and nursing students). Interns and volunteers are excluded.

and repeated measurements). All residents with at least one observation with complete values for fixed effects and outcome were included in each model.

To test for selection bias that could potentially be cause by the ongoing inclusion of new residents, two sensitivity analyses for the primary outcome QoL-AD proxy were performed. The first sensitivity analysis was based on a sample of all residents included at baseline without ongoing inclusion of new residents at T1 and T2. In the second sensitivity analysis, only residents that completed questionnaires at all three time points were included. In further secondary analyses, the same models were fitted using the subscores as the dependent variables. Adjustment for multiple testing was not performed, and these results are interpreted with this limitation taken into account.

The level of significance was 5%. All tests and confidence intervals were two-sided. The statistical analysis was performed using the statistical software packages SAS 9.4 and R version 3.0.1 (R Core Team, 2014).

Results

Participants

All nine NHs completed the study. From these units, 217 residents were eligible at baseline, and 234 residents (Table 1) were incorporated into

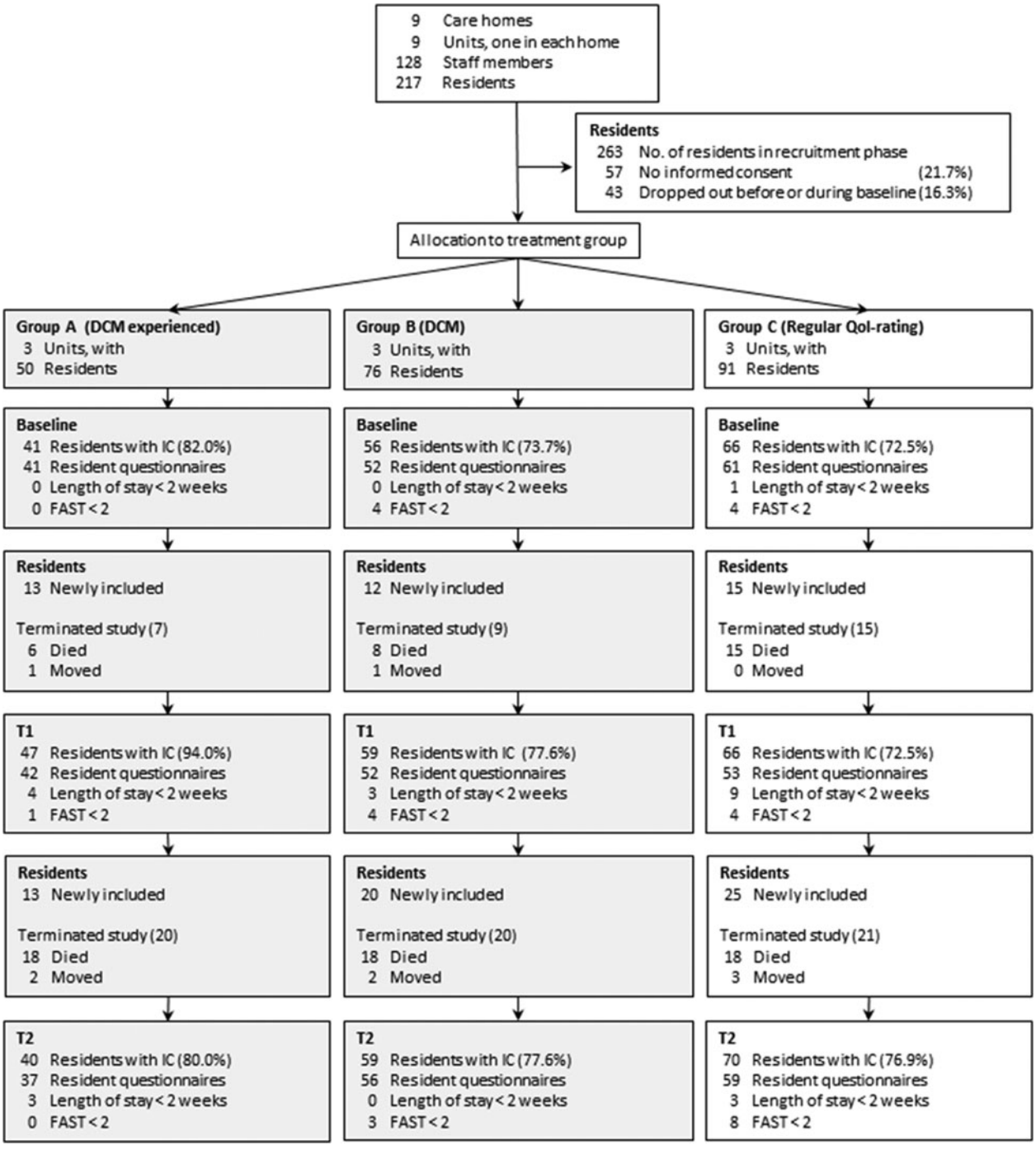

Figure 1. Flowchart with detailed numbers of residents with dementia.

the analysis set, thereby resulting in a total of 453 completed resident questionnaires for all of the time points (Figure 1). Ninety-two residents with dementia did not complete the study, and an additional 98 new residents were included during the study.

Table 1 presents the baseline characteristics of the residents and the units (staff characteristics and dementia milieu). The baseline characteristics of the residents were generally comparable among the groups, with the exception of pain medication use ($p = 0.001$). Differences between staff characteristics and the dementia-friendliness of the unit milieu were only descriptively computed because of the small number of units.

Intervention effects

The overall effects on the primary outcome, QoL of PwD, measured with the QoL-AD proxy for

treatment (two interventions and one control), time (T0, T1, and T2) and interaction between treatment and time were not significant (covariate adjusted model in Table 2: $p_g = 0.23$, $p_t = 0.37$, $p_{gt} = 0.10$).

The results were similar to the model without covariate adjustment (Table 3) and those of sensitivity analyses based on both a sample of all residents included at baseline without ongoing inclusion and a sample of residents with questionnaires at all three time points (see Table S1, available as supplementary material attached to the electronic version of this paper at www.journals.cambridge.org/jid_IPG).

The estimated least square means of the QoL-AD proxy decreased between baseline (T0) and the 18-month follow-up (T2) from 32.54 (confidence interval, hereafter CI: 29.36–35.72) to 31.32 (CI: 28.15–34.48) in group A and from 33.62 (CI: 30.55–36.68) to 27.60 (CI: 24.51–30.69) in group B. For group C, the QoL-AD proxy scores increased from 30.50 (CI: 27.47–33.52) at baseline to 32.54 (CI: 29.44–35.64) after 18 months (Figure 2).

Overall, there were no significant effects of time, treatment, or time * treatment on challenging behavior (covariate adjusted model in Table 2: $p_g = 0.80$, $p_t = 0.09$, $p_{gt} = 0.21$). The estimates increased over time for group A (17.76 (CI: 9.11–26.41) to 21.12 (CI: 12.41–29.83)), and they decreased for group B (25.27 (CI: 16.88–33.66) to 16.95 (CI: 8.57–25.33)) and group C (23.67 (CI: 15.42–31.92) to 13.22 (CI: 4.88–21.56); Figure 3). A significant decrease in apathetic behavior between T0 and T2 ($p_t = 0.03$) in all groups was identified (Table 4). The results of the sensitivity analysis after square root transformation of NPI-NH were similar in terms of the time course and statistical tests (data not shown).

The QoL of residents with dementia measured with QUALIDEM also exhibited no significant effects for treatment, time, and their interaction (Table 2). With decreases in QoL for group A (71.82 (CI: 66.10–77.55) to 69.83 (CI: 64.02–75.64)) and group B (74.77 (CI: 69.16–80.38) to 66.18 (CI: 60.46–71.90)) and an increase for group C (70.36 (CI: 64.83–75.89) to 72.91 (CI: 67.29–78.52)), the development of the measured QoL corresponds to the results of the QoL-AD proxy (Figure 4). Significant effects were identified for the QUALIDEM subscales "positive effect" ($p_{gt} = 0.01$) and "positive self-image" ($p_{gt} = 0.02$). Between baseline and T2, the QoL domain "positive effect" decreased in the DCM groups A and B and increased in group C. The positive self-image of the residents increased in groups A and C and decreased in group B (Table 4).

Discussion

The results of our pragmatic quasi-experimental trial regarding the change in QoL (QoL-AD proxy) indicated no significant differences either between the two DCM intervention groups and the control group or between the two DCM intervention groups. The primary outcome, QoL, exhibited a non-expected and non-significant negative trend for both DCM groups. For the secondary outcomes, challenging behavior (NPI-NH) and QoL (QUALIDEM), non-significant differences between groups were observed. The results based on the QUALIDEM data confirmed the results for the primary outcome, whereas the challenging behavior data exhibited non-significant reductions in groups B and C and a slight increase in group A. For the subscales of these secondary outcome measures, a significant decrease in apathetic behavior (NPI-NH) overall and an increase in positive self-image (QUALIDEM) except for group B were observed. In contrast, the QoL dimension of "positive effect" (QUALIDEM) exhibited a decrease over time for both DCM groups. Based on both the heterogeneous results for the outcome measures and their subscales and the methodological difficulties associated with multiple testing, these results should be interpreted cautiously.

The three previous trials regarding the effectiveness of the DCM method yielded heterogeneous results. Our results can be best compared with the design and results of van de Ven *et al.* (2013). In this pragmatic trial, no positive effects were observed for the DCM method compared with usual care in terms of challenging behavior and QoL. For the primary outcome and the secondary outcome "challenging behavior," a negative non-significant trend for the DCM group compared with the control group was observed. Both studies were based on pragmatic approaches in terms of the lack of inclusion criteria for the participating nursing units. Moreover, the DCM method was performed by trained members of the care staff rather than members of the research teams. No significant effect after eight months of DCM was observed in the study by van de Ven (2013), and we observed no effect after 18 months of intervention. The two other aforementioned studies regarding the effectiveness of DCM, Rokstad *et al.* (2013) and Chenoweth *et al.* (2009), used a more explanatory approach. In the Norwegian trial, DCM was performed by either researchers or trained care staff, depending on the DCM component, and no inclusion criteria for NHs were defined. In this study, no effect was observed for the primary outcome "agitation," but

Table 2. Intervention effects on residents based on an intention-to-treat analysis (adjusted for the following covariates: PSMS, FAST, pain medication use, and NPI-NH total score)

OVERALL $N = 234$ RESIDENTS (1 * 96; 2 * 57; 3 * 81)[a]	BASELINE, T0 ($N = 154$) ESTIMATED SCORE [95% CI]	FOLLOW-UP AFTER 6 MONTHS, T1 ($N = 147$) ESTIMATED SCORE [95% CI]	FOLLOW-UP AFTER 18 MONTHS, T2 ($N = 152$) ESTIMATED SCORE [95% CI]
QoL-AD total score (13–52, primary outcome) $n = 218$ residents (1 * 90; 2 * 68; 3 * 60)[a]			$p_g = 0.23$, $p_t = 0.37$, $p_{gt} = 0.10$
Group A (DCM experienced)	32.54 [29.36–35.72]	32.83 [29.66–35.99]	31.32 [28.15–34.48]
Group B (DCM)	33.62 [30.55–36.68]	29.23 [26.17–32.29]	27.60 [24.51–30.69]
Group C (Regular QoL-rating)	30.50 [27.47–33.52]	32.58 [29.52–35.64]	32.54 [29.44–35.64]
Covariates for adjustment: mean; regression parameter [95% CI]		PSMS-score[b]	19.96; −0.22 [−0.34–−0.09]
		Prevalence FAST = 7[b]	35%; −3.47 [−4.87–−2.08]
		Pain medication use dichotomized[b]	30%; −1.68 [−2.95–−0.40]
		NPI-NH total score[b]	20.09; −0.09 [−0.12–−0.06]
NPI-NH total score (0–144, secondary outcome) $n = 226$ residents (1 * 90; 2 * 68; 3 * 68)[a]			$p_g = 0.80$, $p_t = 0.09$, $p_{gt} = 0.21$
Group A (DCM experienced)	17.76 [9.11–26.41]	24.83 [16.11–33.55]	21.12 [12.41–29.83]
Group B (DCM)	25.27 [16.88–33.66]	22.36 [13.95–30.77]	16.95 [8.57–25.33]
Group C (Regular QoL-rating)	23.67 [15.42–31.92]	19.62 [11.25–28.00]	13.22 [4.88–21.56]
Covariates for adjustment: mean; regression parameter [95% CI]		PSMS-score[b]	20.09; 0.74 [0.34–1.14]
		Prevalence FAST = 7	36%; −0.94 [−5.29–3.41]
		Pain medication use dichotomized	30%; −1.21 [−5.39–2.96]
QUALIDEM total score (0–100, secondary outcome) $n = 219$ residents (1 * 87; 2 * 73; 3 * 59)[a]			$p_g = 0.85$, $p_t = 0.20$, $p_{gt} = 0.15$
Group A (DCM experienced)	71.82 [66.10–77.55]	68.17 [62.36–73.98]	69.83 [64.02–75.64]
Group B (DCM)	74.77 [69.16–80.38]	69.37 [63.76–74.98]	66.18 [60.46–71.90]
Group C (Regular QoL-rating)	70.36 [64.83–75.89]	71.12 [65.53–76.72]	72.91 [67.29–78.52]
Covariates for adjustment: mean; regression parameter [95% CI]		PSMS-score[b]	20.02; −0.31 [−0.56–−0.07]
		Prevalence FAST = 7	36%; −1.57 [−4.17–1.02]
		Pain medication use dichotomized	31%; −0.16 [−2.63–2.32]
		NPI-NH total score[b]	20.16; −0.46 [−0.52–−0.40]

[a]No. of times per resident.
[b]Significant; $p < 0.05$ for covariate.
CI 95% = Confidence interval 95%.
Estimated score = estimated least square means.
p_g = main effect of the intervention.
p_t = main effect of time (at three time points).
p_{gt} = main effect of interaction between group and time.
QoL-AD = Quality of life Alzheimer's disease scale.
NPI-NH = Neuropsychiatric Inventory – Nursing Home version.

Table 3. Intervention effects on residents based on an intention-to-treat analysis (not adjusted for covariates)

Overall $N = 234$ residents (1 * 96; 2 * 57; 3 * 81)[a]	Baseline, T0 ($N = 154$) Estimated score [95% CI]	Follow-up after 6 months, T1 ($N = 147$) estimated score [95% CI]	Follow-up after 18 months, T2 ($N = 152$) estimated score [95% CI]
QoL-AD total score (13–52, primary outcome) $n = 226$ residents (1 * 94; 2 * 62; 3 * 70)			$p_g = 0.45$, $p_t = 0.34$, $p_{gt} = 0.22$
Group A (DCM experienced)	32.73 [28.74–36.71]	32.50 [28.57–36.44]	30.90 [26.92–34.87]
Group B (DCM)	33.36 [29.51–37.21]	29.63 [25.79–33.47]	27.28 [23.43–31.14]
Group C (Regular QoL-rating)	30.87 [27.05–34.69]	33.13 [29.28–36.97]	32.80 [28.93–36.67]
NPI-NH total score (0–144, secondary outcome) $n = 228$ residents (1 * 92; 2 * 66; 3 * 70)[a]			$p_g = 0.71$ $p_t = 0.23$ $p_{gt} = 0.23$
Group A (DCM experienced)	16.94 [7.89–25.99]	24.89 [15.76–34.03]	21.21 [12.09–30.32]
Group B (DCM)	25.35 [16.57–34.14]	22.15 [13.34–30.97]	17.81 [9.07–26.56]
Group C (Regular QoL-rating)	22.49 [13.83–31.15]	18.13 [9.37–26.90]	13.80 [5.07–22.52]
QUALIDEM total score (0–100, secondary outcome) $n = 226$ residents (1 * 92; 2 * 63; 3 * 71)[a]			$p_g = 0.46$, $p_t = 0.15$, $p_{gt} = 0.06$
Group A (DCM experienced)	73.50 [67.11–79.88]	64.96 [58.54–71.38]	68.44 [61.91–74.96]
Group B (DCM)	73.08 [66.93–79.23]	68.75 [62.61–74.90]	66.31 [60.07–72.55]
Group C (Regular QoL-rating)	70.19 [64.14–76.25]	71.79 [65.66–77.92]	76.37 [70.24–82.51]

[a]No. of times per resident.
CI 95% = Confidence interval 95%.
Estimated score = estimated least square means.
p_g = main effect of the intervention.
p_t = main effect of time (at three time points).
p_{gt} = main effect of interaction between group and time.
QoL-AD = Quality of life Alzheimer's disease scale.
NPI-NH = Neuropsychiatric Inventory – Nursing Home version.

positive effects were observed for the secondary outcomes "challenging behavior (overall)" and "QoL" (Rokstad *et al.*, 2013). In Chenoweth *et al.* (2009), the first effectiveness study, the DCM method was performed by two researchers in collaboration with the care staff, and the participating nursing units and the residents with dementia were selected specifically. Thus, the effectiveness of the DCM method was tested under near-ideal conditions (van de Ven *et al.*, 2013). In this trial, a positive effect with regard to the primary outcome "agitation" was observed. No effects for the secondary outcomes "challenging behavior (overall)" and "QoL" were observed (Chenoweth *et al.*, 2009).

Summarizing these results, the effect of DCM seems to depend on the level of support provided by members of the research team during the implementation, which results in no variation of the implementation among the nursing units. The usage of selection criteria for participating NHs may be another explanation for the differences in the results. In the study of Chenoweth *et al.* (2009), only units with task-focused systems, rather than PCC systems, were included. Moreover, the included residents exhibited persistent challenging behavior. In contrast, our study and the study by van de Ven *et al.* (2013) did not employ such selection criteria. In our trial, this decision resulted in a heterogeneous sample in terms of the living arrangements and care concepts of the nursing units. It can be assumed that a heterogeneous NH sample requires varying support for the intervention implementation, which suggests a need for external support, as noted above.

However, it is unsatisfactory that over a period of 18 months, it was not possible to demonstrate a difference between the intervention and control groups – i.e. the experienced and inexperienced DCM groups, respectively – for such a resource-consuming intervention. Neither our study nor two other studies (Rokstad *et al.*, 2013; van de Ven *et al.*, 2013) used the persistence of challenging behavior

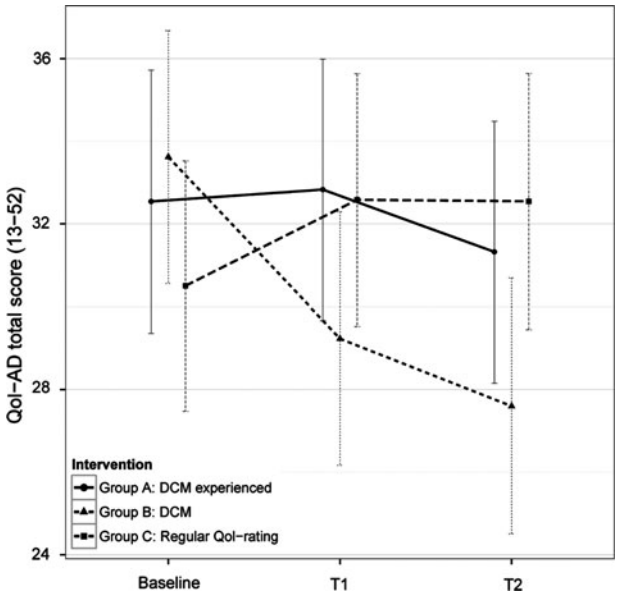

Figure 2. Estimated least square means [CI = 95%] for QoL (QoL-AD-proxy) of residents, adjusted for covariates.

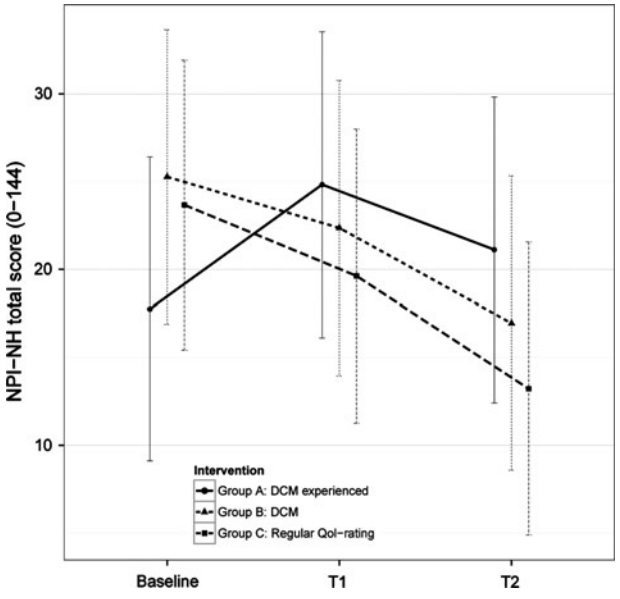

Figure 3. Estimated least square means [CI = 95%] for challenging behavior (NPI-NH) of residents, adjusted for covariates.

as an inclusion criterion for PwD. It is possible that use of such an inclusion criterion increases the likelihood of identifying a positive effect for DCM because the study population is consequently more homogeneous. However, the aim of DCM is PCC, not only reduction of challenging behavior. Thus, because PCC targets the whole person, it should be possible to identify positive effects for prevention of challenging behavior in addition to maintenance and enhancement of the QoL of PwD.

We used QoL as a primary outcome in our trial, whereas the three previous studies used agitation and QoL as primary and secondary outcome measures, respectively. We chose QoL as the primary outcome because the aim of PCC, as defined by DCM, is the enhancement of well-being, which can be observed either as a major dimension of QoL or as a synonym for QoL. We think that the objective of a comprehensive care approach such as PCC must be maintenance and promotion of the QoL of PwD. However, a recent systematic review of the effectiveness of non-pharmacological interventions concluded that only group cognitive stimulation therapy is an effective intervention for improving the QoL of PwD in NHs (Cooper *et al.*, 2012). No effects or only negative effects were observed for other psychosocial and complex interventions, such as DCM (Cooper *et al.*, 2012). This finding may be due to the challenge of measuring QoL in PwD (Scholzel-Dorenbos *et al.*, 2007; Dichter *et al.*, 2013) and the possibility that an improvement in QoL requires more time for manifestation (Cooper *et al.*, 2012). We used the two QoL instruments that are recommended as outcome measures for psychosocial intervention research in dementia care, QoL-AD proxy, and QUALIDEM (Moniz-Cook *et al.*, 2008). Despite the long intervention period of 18 months, it was not possible to demonstrate a significant intervention effect regarding dementia-specific QoL. Nevertheless, there are further unanswered methodological questions regarding the reliability (Dichter *et al.*, 2014), responsiveness (Perales *et al.*, 2013), and the perspective of proxy measures to assess dementia-specific QoL. In particular, the effect of different proxy perspectives on QoL measures such as the QoL-AD proxy has to be investigated in further research. We used the QoL-AD based on a proxy–proxy perspective, which means that the care staff rated the QoL of the resident from the proxy perspective. This perspective differs more from a QoL self-rating performed by a resident than it does from a patient–proxy perspective. In the latter perspective, a proxy assesses the QoL of a resident as the proxy thinks that the resident would rate him or herself (Pickard and Knight, 2005). To date, the patient–proxy perspective has not been regularly used. However, it is possible to use this proxy, and it would potentially yield a QoL rating that is more consistent with the resident's perspective. The advantage of a patient–proxy perspective for evaluation of psychosocial interventions, such DCM is stronger consideration of the individual resident's perspective.

These methodological questions also apply to the assessment of challenging behavior. Beyond the aforementioned difficulties, the effect of the

Table 4. Intervention effects on residents based on an intention-to-treat analysis (adjusted for the following covariates: PSMS, FAST, and pain medication use)

	BASELINE, T0 ($N = 154$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 6 MONTHS, T1 ($N = 147$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 18 MONTHS, T2 ($N = 152$) MEAN SCORE [95% CI]
QoL-AD total score (13–52, primary outcome)			$p_g = 0.23$, $p_t = 0.37$, $p_{gt} = 0.09$
Group A (DCM experienced)	32.54 [29.36–35.72]	32.83 [29.66–35.99]	31.32 [28.15–34.48]
Group B (DCM)	33.62 [30.55–36.68]	29.23 [26.17–32.29]	27.60 [24.51–30.69]
Group C (Regular QoL-rating)	30.50 [27.47–33.52]	32.58 [29.52–35.64]	32.54 [29.44–35.64]
NPI-NH total score (0–144, secondary outcome)			$p_g = 0.80$ $p_t = 0.09$ $p_{gt} = 0.21$
Group A (DCM experienced)	17.76 [9.11–26.41]	24.83 [16.11–33.55]	21.12 [12.41–29.83]
Group B (DCM)	25.27 [16.88–33.66]	22.36 [13.95–30.77]	16.95 [8.57–25.33]
Group C (Regular QoL-rating)	23.67 [15.42–31.92]	19.62 [11.25–28.00]	13.22 [4.88–21.56]
NPI-NH severity score (F * S) for the subscale of delusions (0–12)			$p_g = 0.92$, $p_t = 0.34$, $p_{gt} = 0.39$
Group A (DCM experienced)	0.84 [−0.15–1.83]	1.69 [0.71–2.67]	1.05 [0.04–2.06]
Group B (DCM)	1.49 [0.57–2.42]	0.77 [−0.15–1.69]	0.85 [−0.07–1.76]
Group C (Regular QoL-rating)	1.28 [0.39–2.17]	1.41 [0.49–2.33]	0.48 [−0.43–1.38]
NPI-NH severity score (F * S) for the subscale of hallucinations (0–12)			$p_g = 0.25$, $p_t = 0.27$, $p_{gt} = 0.24$
Group A (DCM experienced)	0.48 [−0.14–1.11]	0.85 [0.23–1.47]	0.71 [0.07–1.36]
Group B (DCM)	1.24 [0.66–1.81]	0.66 [0.09–1.24]	0.73 [0.00–1.3]
Group C (Regular QoL-rating)	0.70 [0.15–1.26]	0.52 [−0.06–1.09]	−0.06 [−0.63–0.51]
NPI-NH severity score (F * S) for the subscale of agitation/aggression (0–12)			$p_g = 0.36$, $p_t = 0.37$, $p_{gt} = 1.00$
Group A (DCM experienced)	3.55 [2.02–5.08]	3.79 [2.26–5.33]	3.22 [1.66–4.77]
Group B (DCM)	3.08 [1.61–4.54]	3.35 [1.89–4.81]	2.62 [1.16–4.08]
Group C (Regular QoL-rating)	2.39 [0.96–3.83]	2.63 [1.17–4.09]	2.01 [0.56–3.46]
NPI-NH severity score (F * S) for the subscale of depressed mood (0–12)			$p_g = 0.96$, $p_t = 0.46$, $p_{gt} = 0.09$
Group A (DCM experienced)	1.99 [0.70–3.27]	2.38 [1.09–3.66]	1.42 [0.11–2.72]
Group B (DCM)	1.79 [0.56–3.02]	1.20 [−0.03–2.43]	2.44 [1.21–3.66]
Group C (Regular QoL-rating)	2.71 [1.51–3.92]	1.85 [0.62–3.07]	1.44 [0.23–2.66]
NPI-NH severity score (F * S) for the subscale of anxiety (0–12)			$p_g = 0.74$, $p_t = 0.92$, $p_{gt} = 0.14$
Group A (DCM experienced)	1.36 [−0.21–2.93]	2.55 [0.98–4.11]	3.12 [1.53–4.71]
Group B (DCM)	3.40 [1.89–4.90]	2.33 [0.83–3.83]	2.03 [0.53–3.53]
Group C (Regular QoL-rating)	2.48 [1.01–3.96]	1.94 [0.44–3.45]	1.55 [0.06–3.05]

Table 4. Continued.

	BASELINE, T0 ($N = 154$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 6 MONTHS, T1 ($N = 147$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 18 MONTHS, T2 ($N = 152$) MEAN SCORE [95% CI]
NPI-NH severity score (F * S) for the subscale of euphoria (0–12)			
			$p_g = 0.95$, $p_t = 0.66$, $p_{gt} = 0.22$
Group A (DCM experienced)	0.14 [−0.41–0.68]	0.37 [−0.17–0.91]	0.59 [0.04–1.14]
Group B (DCM)	0.56 [0.04–1.07]	0.47 [−0.05–0.99]	0.16 [−0.36–0.67]
Group C (Regular QoL-rating)	0.72 [0.21–1.23]	0.43 [−0.09–0.95]	0.21 [−0.31–0.72]
NPI-NH severity score (F * S) for the subscale of apathy (0–12)			
			$p_g = 0.89$, $p_t = 0.03$, $p_{gt} = 0.60$
Group A (DCM experienced)	2.27 [0.83–3.71]	2.89 [1.46–4.33]	1.62 [0.16–3.07]
Group B (DCM)	2.46 [1.08–3.84]	3.39 [2.01–4.77]	1.53 [0.15–2.90]
Group C (Regular QoL-rating)	2.85 [1.49–4.21]	2.16 [0.78–3.54]	1.39 [0.02–2.76]
NPI-NH severity score (F * S) for the subscale of disinhibition (0–12)			
			$p_g = 0.78$, $p_t = 0.77$, $p_{gt} = 0.28$
Group A (DCM experienced)	0.85 [0.12–1.58]	0.92 [0.19–1.64]	1.46 [0.69–2.22]
Group B (DCM)	1.44 [0.79–2.10]	1.58 [0.92–2.24]	0.83 [0.00–1.48]
Group C (Regular QoL-rating)	1.13 [0.53–1.74]	1.24 [0.59–1.89]	0.89 [0.28–1.51]
NPI-NH severity score (F* S) for the subscale of irritability (0–12)			
			$p_g = 0.17$, $p_t = 0.26$, $p_{gt} = 0.14$
Group A (DCM experienced)	1.44 [0.51–2.38]	2.96 [2.04–3.89]	2.38 [1.40–3.36]
Group B (DCM)	1.78 [0.94–2.62]	2.05 [1.22–2.89]	1.13 [0.29–1.97]
Group C (Regular QoL-rating)	1.66 [0.89–2.43]	1.38 [0.00–2.22]	1.58 [0.79–2.37]
NPI-NH severity score (F * S) for the subscale of aberrant motor behavior (0–12)			
			$p_g = 0.89$, $p_t = 0.22$, $p_{gt} = 0.31$
Group A (DCM experienced)	2.08 [0.68–3.49]	2.90 [1.49–4.30]	2.19 [0.76–3.61]
Group B (DCM)	2.93 [1.58–4.28]	2.05 [0.71–3.40]	1.39 [0.05–2.73]
Group C (Regular QoL-rating)	2.70 [1.38–4.02]	1.66 [0.31–3.02]	1.88 [0.54–3.22]
NPI-NH severity score (F * S) for the subscale of night-time behavior (0–12)			
			$p_g = 0.77$, $p_t = 0.38$, $p_{gt} = 0.19$
Group A (DCM experienced)	0.65 [−0.68–1.97]	1.79 [0.46–3.12]	1.70 [0.37–3.04]
Group B (DCM)	2.49 [1.22–3.77]	1.51 [0.24–2.79]	1.21 [0.06–2.48]
Group C (Regular QoL-rating)	1.88 [0.63–3.13]	1.90 [0.63–3.17]	0.42 [−0.84–1.68]
NPI-NH severity score (F * S) for the subscale of eating change (0–12)			
			$p_g = 0.97$, $p_t = 0.09$, $p_{gt} = 0.75$
Group A (DCM experienced)	1.94 [0.42–3.46]	2.05 [0.53–3.57]	1.59 [0.06–3.13]
Group B (DCM)	2.49 [1.02–3.97]	2.15 [0.69–3.62]	1.40 [−0.06–2.86]
Group C (Regular QoL-rating)	2.89 [1.44–4.34]	2.09 [0.62–3.56]	1.04 [−0.42–2.50]

Table 4. Continued.

	BASELINE, T0 ($N = 154$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 6 MONTHS, T1 ($N = 147$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 18 MONTHS, T2 ($N = 152$) MEAN SCORE [95% CI]
QUALIDEM total score (0–100, secondary outcome)			
			$p_g = 0.85$, $p_t = 0.20$, $p_{gt} = 0.15$
Group A (DCM experienced)	71.82 [66.10–77.55]	68.17 [62.36–73.98]	69.83 [64.02–75.64]
Group B (DCM)	74.77 [69.16–80.38]	69.37 [63.76–74.98]	66.18 [60.46–71.90]
Group C (Regular QoL-rating)	70.37 [64.84–75.90]	71.12 [65.53–76.72]	72.91 [67.29–78.52]
QUALIDEM subscale of care relationship (0–100)			
			$p_g = 0.86$, $p_t = 0.11$, $p_{gt} = 0.50$
Group A (DCM experienced)	73.13 [62.55–83.71]	67.34 [56.79–77.90]	78.71 [68.00–89.43]
Group B (DCM)	77.28 [67.02–87.54]	74.59 [64.35–84.83]	74.73 [64.45–85.01]
Group C (Regular QoL-rating)	74.89 [64.76–85.01]	73.26 [63.03–83.50]	79.89 [69.69–90.09]
QUALIDEM subscale of positive affect (0–100)			
			$p_g = 0.92$, $p_t = 0.88$, $p_{gt} = 0.01$
Group A (DCM experienced)	80.98 [71.42–90.53]	69.77 [60.16–79.38]	69.97 [60.30–79.64]
Group B (DCM)	78.06 [68.94–87.18]	70.63 [61.51–79.75]	67.35 [58.20–76.50]
Group C (Regular QoL-rating)	61.99 [53.05–70.94]	75.37 [66.26–84.49]	80.27 [71.18–89.37]
QUALIDEM subscale of negative affect (0–100)			
			$p_g = 0.60$, $p_t = 0.31$, $p_{gt} = 0.45$
Group A (DCM experienced)	74.62 [62.48–86.77]	67.84 [55.72–79.97]	72.84 [60.61–85.07]
Group B (DCM)	75.81 [64.00–87.63]	68.86 [57.04–80.68]	69.11 [57.28–80.93]
Group C (Regular QoL-rating)	73.99 [62.27–85.71]	73.89 [62.09–85.69]	83.12 [71.36–94.89]
QUALIDEM subscale of restless, tense behavior (0–100)			
			$p_g = 0.49$, $p_t = 0.65$, $p_{gt} = 0.42$
Group A (DCM experienced)	65.76 [48.64–82.88]	62.48 [45.37–79.59]	52.18 [34.94–69.41]
Group B (DCM)	59.81 [43.02–76.59]	59.32 [42.54–76.10]	53.02 [36.24–69.80]
Group C (Regular QoL-rating)	67.17 [50.50–83.84]	63.15 [46.38–79.92]	73.62 [56.89–90.35]
QUALIDEM subscale of positive self-image (0–100)[a]			
			$p_g = 0.02$, $p_t = 0.01$, $p_{gt} = 0.02$
Group A (DCM experienced)	70.80 [61.86–79.73]	63.73 [55.28–72.18]	75.90 [66.79–85.01]
Group B (DCM)	83.35 [75.74–90.96]	72.02 [64.79–79.25]	75.24 [67.14–83.33]
Group C (Regular QoL-rating)	76.99 [70.37–83.60]	76.86 [69.78–83.95]	91.46 [84.33–98.60]
QUALIDEM subscale of social relations (0–100)			
			$p_g = 0.40$, $p_t = 0.68$, $p_{gt} = 0.19$
Group A (DCM experienced)	76.67 [67.87–85.46]	67.23 [58.47–76.00]	72.39 [63.49–81.30]
Group B (DCM)	72.82 [64.44–81.29]	70.92 [62.54–79.30]	68.45 [60.00–76.89]
Group C (Regular QoL-rating)	62.06 [53.84–70.28]	66.86 [58.49–75.22]	70.12 [61.82–78.42]

Table 4. Continued.

	BASELINE, T0 ($N = 154$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 6 MONTHS, T1 ($N = 147$) MEAN SCORE [95% CI]	FOLLOW-UP AFTER 18 MONTHS, T2 ($N = 152$) MEAN SCORE [95% CI]
QUALIDEM subscale of social isolation (0–100)			
			$p_g = 0.67$, $p_t = 0.30$, $p_{gt} = 0.63$
Group A (DCM experienced)	75.52 [62.19–88.85]	71.10 [57.78–84.42]	77.45 [64.07–90.83]
Group B (DCM)	76.17 [63.08–89.26]	70.92 [57.84–84.01]	74.13 [61.04–87.22]
Group C (Regular QoL-rating)	75.60 [62.59–88.62]	78.78 [65.70–91.85]	85.04 [71.99–98.10]
QUALIDEM subscale of feeling at home (0–100)			
			$p_g = 0.59$, $p_t = 0.45$, $p_{gt} = 0.11$
Group A (DCM experienced)	74.68 [63.79–85.58]	75.24 [64.52–85.95]	73.59 [62.43–84.74]
Group B (DCM)	85.91 [75.61–96.22]	74.20 [64.05–84.35]	75.18 [64.70–85.65]
Group C (Regular QoL-rating)	77.76 [67.76–87.76]	78.85 [68.74–88.95]	84.64 [74.33–94.94]
QUALIDEM subscale of having something to do (0–100)[a]			
			$p_g = 0.73$, $p_t = 0.80$, $p_{gt} = 0.55$
Group A (DCM experienced)	51.69 [37.70–65.69]	41.24 [27.49–54.98]	44.01 [29.50–58.53]
Group B (DCM)	47.65 [34.59–60.71]	49.16 [36.35–61.96]	42.24 [28.88–55.60]
Group C (Regular QoL-rating)	47.15 [34.64–59.67]	52.52 [39.82–65.23]	51.19 [38.22–64.16]

CI 95% = Confidence interval 95%.
p_g = main effect of the intervention.
p_t = main effect of time (at three time points).
p_{gt} = main effect of interaction between group and time.
QoL-AD = Quality of life Alzheimer's disease scale.
NPI-NH = Neuropsychiatric Inventory – Nursing Home version.
[a]Only the PSMS and pain medication use covariates are included in the model.

intervention on the external assessment is unclear. Thus, it is possible that based on the interventions, the care staff members developed more sensitive observation skills over time, which attenuated any effect between intervention and control. This phenomenon is also possible for the outcome QoL. It is conceivable that the focus of the observations of care staff members correlated with the targets of the intervention and the content of the assessed proxy instruments, which necessitates the need for resident self-ratings as long as possible combined with patient–proxy-based or observation-based measurements as outcomes for the evaluation of psychosocial interventions in dementia care.

Based on the theoretical assumptions and intervention components, negative effects of DCM on resident seem to be unlikely. However, negative effects are possible because of, e.g. invalid DCM observations or incorrect interpretation of the results of DCM observations. Regarding staff outcomes, DCM may have negative effects in terms of, e.g. job satisfaction or burnout, especially under

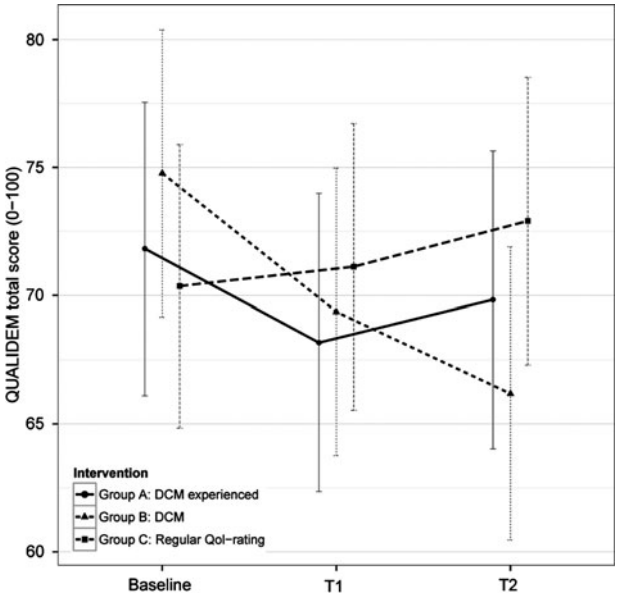

Figure 4. Estimated least square means [CI = 95%] for QoL (QUALIDEM) of residents, adjusted for covariates.

poor institutional conditions at the NH level (e.g. staff shortages or poor leadership). In addition, negative staff outcomes could also affect proxy-based resident outcomes. The results of the process evaluation of the Leben-QD II trial (Halek *et al.*, 2013) will help identify possible negative effects of DCM based on interviews with the participating care staff. In addition, the results of the process evaluation will provide information about specific facilitators and barriers for DCM. This information will help to design future DCM effectiveness trials.

Study limitations

Because of the small number of nursing units in each intervention group ($n = 3$) and thus low power, type II error is possible. A power estimate determined via a two-sided cluster-adjusted *t*-test (Donner and Klar, 2000) for the primary outcome QoL-AD-proxy (comparing two groups) can serve as a (possibly conservative) estimate for the power. Assuming a mean cluster size of 17 at the last time point (T2), a common standard deviation of 6, and an intra-cluster correlation coefficient (ICCC) of 0.2, a QoL-AD difference of $\pm$ 4 could be detected with a power of 90% if five clusters could be recruited per group (Cronbach's $\alpha = 5\%$) (Donner and Klar, 2000). The above assumptions regarding the parameters correspond to the estimates in this trial except that the ICCC estimates were heterogeneous in the small groups with three clusters per group. A difference of $\pm$ 3 is considered clinically relevant (Logsdon *et al.*, 2002). The power estimate suggests that more than five NH units per group would be necessary to detect a clinical relevant difference in the positive or the negative direction (two-sided test) between two groups at the last follow up. Furthermore, the 95% confidence intervals of the QoL estimates are large because of the low power. A more sophisticated power estimate that considers the complete study design with three groups and three time points would require specific assumptions regarding the alternative (which was not defined in the study plan), would be difficult to achieve based on simulations, and would add only a small amount of very specific information.

Furthermore, we were not able to randomly allocate the clusters to the intervention groups. Blinding of the proxy-rating care staff members was not possible because they were included in the intervention while they were functioning as regular staff members on the participating nursing units. The proxy-rating nurses were also directly included in the intervention as trained DCM basic users or team members who were involved in the DCM feedback meetings or action planning. Their direct involvement in the intervention might have led to an overestimation of the QoL of PwD. Conversely, the proxy-rated QoL ratings of the care staff could have been negatively influenced by the challenging behavior of PwD, nurses' burnout and their life satisfaction (Graske *et al.*, 2014a). Because of the ongoing inclusion of residents, selection bias (e.g. based on different lengths of the exposure to the DCM intervention) may have occurred. However, sensitivity analyses that both included and excluded the residents included during the study yielded the same results.

Regarding the motivation of the NH providers to participate in this study, the included nursing units do not appear to be representative of German NH units in general. The cases that were lost to follow-up might have resulted in some bias. The analyses based on mixed models used all of the available information about the participants, even if their data were not documented at all three time points. Nevertheless, we stress that our results are in agreement with those of a large pragmatic cRCT that studied the effectiveness of DCM (van de Ven *et al.*, 2013). The main strength of our study is the long evaluation period, which extended over 18 months in a real-life setting. Thus, our study contributes to the body of knowledge regarding the effectiveness of DCM. Another strength is the small number of missing values for the participating residents at each time point, which is the result of a standardized data collection procedure that used trained student assistants.

Conclusions

Despite our study's limitations, the results of our study demonstrate that the DCM method had no statistically significant effect on the QoL or challenging behavior in PwD in the real-life setting of the participating German NHs. Compared with the previous trials (Chenoweth *et al.*, 2009; Rokstad *et al.*, 2013), it could be hypothesized that the likelihood of the identification of a DCM effect increases if the application of the DCM components is supported by members of the research team as external experts. This support must be embedded in an implementation strategy that targets the heterogeneity of nursing units, based on an analysis of the settings, stakeholders, and facilitators and the barriers to the implementation of PCC. Furthermore, our results indicate no effect of DCM after an extended evaluation period of 18 months, regardless of whether the units had prior DCM experience. For further research, use of patient–proxy perspective or an observational-based outcome is recommended for the evaluation of psychosocial interventions. Moreover, possible negative effects of DCM on

resident outcomes must be considered during the design of future DCM trials.

Conflict of interest

The study is partially funded by the Public Welfare Foundation, North Rhine-Westphalia (Stiftung Wohlfahrtspflege Nordrhein-Westfalen), and the Knights of St John of Jerusalem Nursing Home Section West (Johanniter Seniorenhäuser GmbH, Regionalzentrum West). The Knights of St John of Jerusalem Nursing Home Section West is also the provider that operates the institutions included in the study. The Knights of St John of Jerusalem Nursing Home Section West is a well-established welfare organization with a charitable status. The Public Welfare Foundation North Rhine-Westphalia is a public foundation that supports projects in the field of social welfare with public funding. Neither organization has competing interests. Most of the authors are employees of DZNE and declare that they have no competing interests. CR is Strategic Lead for DCM in Germany. In this role, neither she nor the German Strategic Partner of the Bradford Dementia Group had any influence on planning or conducting the study. BH is self-employed with mediStatistica and received a fee for statistical consulting and analyses. He has no competing interests.

Description of authors' roles

Study Design:	MND, MH, CR, TQ, SB
Data Collection:	MND, CR, TQ, CGGS
Data Analysis:	MND, DT, BH, MH
First Draft of the Manuscript:	MND, MH, DT, BH,
Manuscript Preparation:	MND, MH, CR, TQ, CGGS, SB, DT, BH

Acknowledgments

We would like to thank Stefan Ortner for his contribution as an in-house DCM trainer during the trial. We also thank the participating people with dementia, nurses, and managers of the nursing home teams and study assistants for their contributions.

References

Aged Care Branch (2011). *Dementia-Friendly Environments. A Guide for Residential care*. Canberra: Department of Health, Australia.

Bartholomeyczik, S. *et al.* (2007). *Guidelines for Handling Challenging behavior in People with Dementia Living in Nursing Homes*. Berlin: Bundesministerium für Gesundheit (BMG).

Beerens, H. C., Zwakhalen, S. M., Verbeek, H., Ruwaard, D. and Hamers, J. P. (2013). Factors associated with quality of life of people with dementia in long-term care facilities: a systematic review. *International Journal of Nursing Studies*, 50, 1259–1270.

Brooker, D. (2004). What is person-centred care in dementia? *Reviews in Clinical Gerontology*, 13, 215–222.

Brooker, D. and Surr, C. (2007). *Dementia Care Mapping: Principles and Practice*, Bradford: University of Bradford.

Brooker, D. J. and Surr, C. (2006). Dementia care mapping (DCM): initial validation of DCM 8 in UK field trials. *International Journal of Geriatric Psychiatry*, 21, 1018–1025.

Brown, H. and Prescott, R. (2006). *Applied Mixed Models in Medicine. Statistics in Practice*. Hoboken: John Wiley & Sons.

BSI - British Standards Institution and University of Bradford (2010). *PAS 800: 2010. Use of Dementia Care Mapping for Improved Person-Centred care in a care Provider Organisation - Guide*. London: BSI-British Standards Institution.

Chenoweth, L. *et al.* (2009). Caring for aged dementia care resident study (CADRES) of person-centred care, dementia-care mapping, and usual care in dementia: a cluster-randomised trial. *Lancet Neurology*, 8, 317–325.

Cohen-Mansfield, J., Thein, K., Marx, M.S., Dakheel-Ali, M. and Freedman, L. (2012). Efficacy of non-pharmacologic interventions for agitation in advanced dementia: a randomized, placebo-controlled trial. *The Journal of Clinical Psychiatry*, 73, 1255–1261.

Cooper, C. *et al.* (2012). Systematic review of the effectiveness of non-pharmacological interventions to improve quality of life of people with dementia. *International Psychogeriatrics/IPA*, 24, 856–870.

Dichter, M., Bartholomeyczik, S., Nordheim, J., Achterberg, W. and Halek, M. (2011). Validity, reliability, and feasibility of a quality of life questionnaire for people with dementia. *Zeitschrift für Gerontologie und Geriatrie*, 44, 405–410.

Dichter, M. N., Halek, M., Dortman, O., Meyer, G. and Bartholomeyczik, S. (2013). Measuring the quality of life of people with dementia in nursing homes in Germany - the Study Protocol for the Qol-Dem Project [Die Erfassung der Lebensqualität von Menschen mit Demenz in stationären Altenpflegeeinrichtungen in Deutschland - Studienprotokoll des Qol-Dem Projektes]. *GMS Psychosoc Med*, 10, Doc07. doi: 10.3205/psm000097.

Dichter, M. N., Schwab, C. G., Meyer, G., Bartholomeyczik, S., Dortmann, O. and Halek, M. (2014). Measuring the quality of life in mild to very severe dementia: testing the inter-rater and intra-rater reliability of the German version of the QUALIDEM. *International Psychogeriatrics/IPA*, 26, 825–836.

Dichter, M. N. *et al.* (2013a). Scalability and internal consistency of the German version of the dementia-specific quality of life instrument QUALIDEM in nursing homes – a secondary data analysis. *Health and Quality of Life Outcomes*, 11, 91.

Donner, A. and Klar, N. (2000). *Design and Analysis of Cluster Randomization Trials in Health Research*. London: Arnold.

Ettema, T. P., Droes, R. M., de Lange, J., Mellenbergh, G. J. and Ribbe, M. W. (2007). QUALIDEM: development and evaluation of a dementia specific quality of life instrument. Scalability, reliability, and internal structure. *International Journal of Geriatric Psychiatry*, 22, 549–556.

Graske, J., Meyer, S. and Wolf-Ostermann, K. (2014a). Quality of life ratings in dementia care – a cross-sectional study to identify factors associated with proxy-ratings. *Health Qual Life Outcomes (HQLO)*, 12, 177.

Graske, J., Verbeek, H., Gellert, P., Fischer, T., Kuhlmey, A. and Wolf-Ostermann, K. (2014b). How to measure quality of life in shared-housing arrangements? A comparison of dementia-specific instruments. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, care, and Rehabilitation*, 23, 549–559.

Halek, M., Dichter, M. N., Quasdorf, T., Riesner, C. and Bartholomeyczik, S. (2013). The effects of dementia care mapping on nursing home residents' quality of life and staff attitudes: design of the quasi-experimental study Leben-QD II. *BMC Geriatrics*, 13, 53.

Kitwood, T. (1997). *Dementia Reconsidered: The Person Comes first*. Berkshire, UK: Open University Press.

Lawton, M. P. and Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *The Gerontologist*, 9, 179–186.

Logsdon, R. G., Gibbons, L. E., McCurry, S. M. and Teri, L. (1999). Quality of life in Alzheimer's disease: patient and caregiver reports. *Journal of Mental Health and Aging*, 5, 21–32.

Logsdon, R. G., Gibbons, L. E., McCurry, S. M. and Teri, L. (2002). Assessing quality of life in older adults with cognitive impairment. *Psychosomatic Medicine*, 64, 510–519.

MDS - Medizinicher Dienst des Spitzenverbandes Bund der Krankenkassen (2009). Quality assurance criteria of the German medical advisory service of the statutory health insurance for institutional care. Essen.

Meleis, A. I., Sawyer, L. M., Im, E. O., Hilfinger Messias, D. K. and Schumacher, K. (2000). Experiencing transitions: an emerging middle-range theory. *ANS. Advances in Nursing Science*, 23, 12–28.

Moniz-Cook, E. *et al.* (2008). A European consensus on outcome measures for psychosocial intervention research in dementia care. *Aging & Mental Health*, 12, 14–29.

Moyle, W., Mcallister, M., Venturato, L. and Adams, T. (2007). Quality of life and dementia. *Dementia*, 6, 175–191.

NICE (2006). Dementia. Supporting people with dementia and their carers in health and social care. *NICE Clinical Guideline Number 42*. London: National Insitute for Health and Clinical Excellence.

O'Rourke, H. M., Fraser, K. D. and Duggleby, W. (2015). Does the quality of life construct as illustrated in quantitative measurement tools reflect the perspective of people with dementia? *Journal of Advanced Nursing*. doi: 10.1111/jan.12667.

Perales, J., Cosco, T. D., Stephan, B. C., Haro, J. M. and Brayne, C. (2013). Health-related quality-of-life instruments for Alzheimer's disease and mixed dementia. *International Psychogeriatrics/IPA*, 25, 691–706.

Pickard, A. S. and Knight, S. J. (2005). Proxy evaluation of health-related quality of life: a conceptual framework for understanding multiple proxy perspectives. *Medical Care*, 43, 493–499.

R Core Team (2014). *R: A Language and Environmental for Statistical Computing*. Vienna, Austria: Foundation for Statistical Computing. Available at: http://www.R-project.org; last accessed 11 May 2014.

Reisberg, B. (2007). Global measures: utility in defining and measuring treatment response in dementia. *International Psychogeriatrics/IPA*, 19, 421–456.

Rokstad, A. M., Rosvik, J., Kirkevold, O., Selbaek, G., Saltyte Benth, J. and Engedal, K. (2013). The effect of person-centred dementia care to prevent agitation and other neuropsychiatric symptoms and enhance quality of life in nursing home patients: a 10-month randomized controlled trial. *Dementia and Geriatric Cognitive Disorders*, 36, 340–353.

Rosvik, J., Kirkevold, M., Engedal, K., Brooker, D. and Kirkevold, O. (2011). A model for using the VIPS framework for person-centred care for persons with dementia in nursing homes: a qualitative evaluative study. *International Journal of Older People Nursing*, 6, 227–236.

Scholzel-Dorenbos, C. J., van der Steen, M. J., Engels, L. K. and Olde Rikkert, M. G. (2007). Assessment of quality of life as outcome in dementia and MCI intervention trials: a systematic review. *Alzheimer Disease & Associated Disorders*, 21, 172–178.

van de Ven, G. *et al.* (2013). Effects of dementia-care mapping on residents and staff of care homes: a pragmatic cluster-randomised controlled trial. *PloS one*, 8, e67325.

Vernooij-Dassen, M., Vasse, E., Zuidema, S., Cohen-Mansfield, J. and Moyle, W. (2010). Psychosocial interventions for dementia patients in long-term care. *International Psychogeriatrics/IPA*, 22, 1121–1128.

Wood, S. *et al.* (2000). The use of the neuropsychiatric inventory in nursing home residents. Characterization and measurement. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, 8, 75–83.

World Health Organization (WHO) (2003). *Introduction to drug utilization Research*. WHO International Working Group for Drug Statistics Methodology, WHO Collaborating Centre for Drug Statistics Methodology, WHO Collaborating Centre for Drug Utilization Research and Clinical Pharmacological Services. Oslo, Norway.