

ORIGINAL ARTICLE

The MOBID-2 pain scale: Reliability and responsiveness to pain in patients with dementia

B.S. Husebo^{1,2}, R. Ostelo^{3,4}, L.I. Strand¹

1 Department of Global Public Health and Primary Care, Centre for Elderly and Nursing Home Medicine, University of Bergen, Norway

2 Centre for Age-Related Medicine, Stavanger University Hospital, Norway

3 Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands

4 Department of Health Sciences and the EMGO Institute for Health and Care Research, VU University, Amsterdam, The Netherlands

Correspondence

Bettina S. Husebo

E-mail: bettina.husebo@isf.uib.noc

Funding sources

This project was funded by the Norwegian Research Council (Sponsor's Protocol Code: 189439) and the University of Bergen (09/1568).

Conflicts of interest

None declared.

Accepted for publication

5 March 2014

doi:10.1002/ejp.507

Abstract

Background: Mobilization-Observation-Behavior-Intensity-Dementia-2 (MOBID-2) pain scale is a staff-administered pain tool for patients with dementia. This study explores MOBID-2's test-retest reliability, measurement error and responsiveness to change.

Methods: Analyses are based upon data from a cluster randomized trial including 352 patients with advanced dementia from 18 Norwegian nursing homes. Test-retest reliability between baseline and week 2 ($n = 163$), and weeks 2 and 4 ($n = 159$) was examined in patients not expected to change (controls), using intraclass correlation coefficient ($ICC_{2,1}$), standard error of measurement (SEM) and smallest detectable change (SDC). Responsiveness was examined by testing six priori-formulated hypotheses about the association between change scores on MOBID-2 and other outcome measures.

Results: ICCs of the total MOBID-2 scores were 0.81 (0–2 weeks) and 0.85 (2–4 weeks). SEM and SDC were 1.9 and 3.1 (0–2 weeks) and 1.4 and 2.3 (2–4 weeks), respectively. Five out of six hypotheses were confirmed: MOBID-2 discriminated ($p < 0.001$) between change in patients with and without a stepwise protocol for treatment of pain (SPTP). Moderate association ($r = 0.35$) was demonstrated with Cohen-Mansfield Agitation Inventory, and no association with Mini-Mental State Examination, Functional Assessment Staging and Activity of Daily Living. Expected associations between change scores of MOBID-2 and Neuropsychiatric Inventory – Nursing Home version were not confirmed.

Conclusion: The SEM and SDC in connection with the MOBID-2 pain scale indicate that the instrument is responsive to a decrease in pain after a SPTP. Satisfactory test-retest reliability across test periods was demonstrated. Change scores ≥ 3 on total and subscales are clinically relevant and are beyond measurement error.

1. Introduction

Undiagnosed and untreated pain is common in nursing home (NH) patients with advanced dementia. Approximately 40–80% of these individuals are reported to be in pain with a substantial need for adequate pain treatment (Corbett et al., 2012). Under-treatment is espe-

cially marked in the presence of frailty, leading to behavioural disturbances such as agitation (Husebo et al., 2011a,b; Corbett et al., 2012), depression (Cohen-Mansfield and Taylor, 1998), anxiety (Arola et al., 2010) and reduced quality of life (Jakobsson and Hallberg, 2002; Chen et al., 2003). However, assessment and treatment of pain in dementia is complex

What's already known about this topic?

- Identification and monitoring of pain is challenging in patients with dementia.
- Efficacy assessment of pain treatment depends upon the validity of a tool, which should be responsive to change in pain intensity after pain treatment.

What does this study add?

- Following the recommendations by COSMIN panel, MOBID-2 (Mobilization-Observation-Behavior-Intensity-Dementia-2) pain scale was tested regarding reliability, standard error of measurement, smallest detectable change and responsiveness.
- Indications were provided that MOBID-2 is responsive to a decrease in pain intensity after pain treatment over time.

because persons with memory, language and speech deficits are unable to communicate clearly about their pain state, analgesic effects and side effects.

In 2009, the American Geriatric Society recommended a comprehensive, disease-specific assessment to determine appropriate treatment for each individual (American Geriatrics Society Panel on Pharmacological Management of Persistent Pain in Older Persons, 2009). Self-reported pain is often invalid, and pain must therefore be indirectly observed by proxy raters using a validated pain behaviour instrument. More than 30 such tools have been developed, and review articles address the measurement properties of these instruments and feasibility in clinical practice (Hadjistavropoulos et al., 2007; Herr, 2011; Corbett et al., 2012; Husebo et al., 2012). These are important steps towards valid proxy rater *assessment* of pain behaviour in patients with dementia.

The next step to improve *treatment* is the administration of pain medication and evaluation of the effectiveness of this intervention (Corbett et al., 2012). Thus, the development of a practical and responsive tool to capture change in pain intensity after pain treatment is a prerequisite (Terwee et al., 2007). Recently defined by the COSMIN panel, responsiveness is the 'ability of an instrument to detect change over time in the construct to be measured' (Mokkink et al., 2010). When a gold standard is available, correlations between change scores and the area under the receiver operator curve (ROC) are considered (Mokkink et al., 2010). If no gold standard is available, as is the case with patients with dementia, assessment

of responsiveness relies upon testing hypotheses about expected correlations between changes in scores of an instrument and changes in other variables. Although previous studies report promising results regarding the responsiveness of pain behaviour tools for patients with dementia, they are not in line with the methodological requirements regarding a priori-formulated hypotheses, adequate sample size or randomization process (Morello et al., 2007; Cohen-Mansfield and Lipson, 2008; Rat et al., 2011).

The Mobilization-Observation-Behavior-Intensity-Dementia-2 (MOBID-2) pain scale is an observational pain tool for patients with advanced dementia. Earlier studies on measurement properties indicated high to excellent reliability and aspects of validity (Husebo et al., 2007, 2009, 2010), and MOBID-2 proved to be feasible for use in clinical practice (Husebo et al., 2008, 2011b). Responsiveness to change has, however, not been addressed. The aim of this study is to repeat test-retest reliability and to examine the standard error of measurement (SEM) and responsiveness to change in connection with a cluster randomized trial.

2. Methods

2.1 Setting and selection of patients

The study of measurement properties was performed in connection with a cluster randomized controlled trial (RCT) of the efficacy of pain treatment on behavioural disturbances of 352 NH patients with advanced dementia (Husebo et al., 2011b). The patients were from 18 NHs in five municipalities in Western Norway, recruited from October 2009 to June 2010. The patients had lived in the NHs for at least 4 weeks prior to the trial. They had moderate to severe dementia and significant behavioural disturbances. The severity of dementia was assessed by a score of ≤ 19 on the Mini-Mental State Examination (MMSE) scale (range 0–30) (Folstein et al., 1975; Burns et al., 2010). Patients were included independent of painful diagnoses, presumed pain or ongoing pain treatment. Dying patients or patients with acute or unstable psychiatric or medical conditions were not included. The recruitment strategy, patient samples and design of the study have been described elsewhere (Husebo et al., 2011b).

2.2 Ethics

Written informed consent included a description of the study design, benefit and possible side effects of the trial. Since individuals with mild cognitive impairment have an impaired capacity to consent to research (Ayalon, 2009), informed consent was obtained from all patients and all surrogates/caregivers or the authorized legal representatives. In accordance with local law, the study was approved by the Regional Committee for Medical Ethics, Western Norway

(REK-Vest 248.08), by the authorized Institutional Review Board of each participating institution and the Norwegian Medicines Agency (EudraCTnr: 2008-007490-20). In addition, the trial is registered at ClinicalTrials.gov (NCT01021696). The caregivers gave consent to participate as informants.

2.3 Raters

During the enrolment process, two trained research assistants retrieved data pertaining to background variables, which included information about gender, age, medical information (diagnoses and pain diagnoses) and lists of medication taken (Husebo et al., 2011b). The research assistants did not participate in other data collection during the intervention study.

Before starting the enrolment process, the patients' primary caregivers ($n = 53$), usually a licensed practical nurse (LPN) who knew the patient at least for the last 4 weeks, learned how to use measurement instruments during a 2-h education programme. They learned the MOBID-2 procedure by observing video recording of patients with dementia and by bedside observation of the assessment of three patients. Efforts were made to keep the patients and the LPNs uninformed about the study design and type of intervention. This means that the LPNs were blinded to group allocation during assessments of the patients.

A consultant for old age psychiatry, an anaesthetist and pain therapist (B.S.H), one of the research assistants and a senior member of the NH staff from each institution reviewed assessment outcomes and drug prescriptions for each patient after completion of baseline assessment, but before randomization. In addition, they optimized the general treatment; this means, for instance, a patient with symptoms of urinary tract infection received disease-specific diagnostic and treatment.

Registered nurses, responsible for the documentation, prescribing procedures and distribution of the study medication during mealtimes, did not participate in data collection of primary and secondary outcomes. The NH staff were instructed not to discuss management procedures.

2.4 Examination of reliability and responsiveness

2.4.1 Test-retest reliability

Patients included in the reliability study were those randomized to the control group. It was therefore expected that these patients would not systematically improve in their pain state over the time intervals. They were tested with the MOBID-2 pain scale at baseline (week 0) and at weeks 2 and 4. The reliability was assessed by intraclass correlation coefficient ($ICC_{2,1}$) using 2-week intervals (0–2 and 2–4 weeks). Moreover, the SEM and smallest detectable change (SDC) were estimated. Using both the 2-week intervals enabled us to assess the stability of our estimates.

2.4.2 Responsiveness

Following the recommendations by the COSMIN panel (Mokkink et al., 2010), responsiveness was examined by formulating (a priori) and testing six hypotheses in patients expected to be in pain (MOBID-2 pain scale total scores ≥ 3) at baseline.

Hypothesis 1: We hypothesized that pain on the total scale and on the two subscales would decrease more during the intervention period of 8 weeks in the intervention patients who received increased pain medication than in patients who continued with care and treatment as usual (control group).

Hypothesis 2: We hypothesized that change in pain scores on the MOBID-2 pain scale was moderately correlated with change scores on the Cohen-Mansfield Agitation Inventory (CMAI) (Cohen-Mansfield and Libin, 2004). It has previously been shown that untreated pain tends to increase behavioural disturbances in patients with dementia (Chibnall et al., 2005; Kovach et al., 2006; Kovacs et al., 2007). Thus, some behaviours, such as agitation, complaining, pacing or restlessness, may be related to the presence of pain (Husebo et al., 2011a,b). Pain intensity is, however, a different phenomenon than frequency of behavioural disturbances. Therefore, changes in these observed phenomena were not expected to be more than moderately correlated.

Hypothesis 3: We hypothesized that change in pain scores on the MOBID-2 pain scale was moderately correlated with change scores from the Neuropsychiatric Inventory – Nursing Home version (NPI-NH) (Selbaek et al., 2008). Under-treated pain may also increase neuropsychiatric symptoms in patients with dementia (Manfredi et al., 2003; Chibnall et al., 2005; Kovach et al., 2006; Corbett et al., 2012), and there is strong evidence that mood symptoms such as depression, apathy or nocturnal behaviour are related to untreated pain (Husebo et al., 2011a,b, 2013). Pain intensity is, however, a different phenomenon than neuropsychiatric symptoms, and changes in these observed phenomena were not expected to be more than moderately related.

Hypotheses 4–6: We finally hypothesized that change scores of pain on MOBID-2 pain scale were not correlated with change scores on measures of the Barthel index of Activity of Daily Living (ADL) (Mahoney and Barthel, 1965), the MMSE (Folstein et al., 1975) and the Functional Assessment Staging (FAST) (Hughes et al., 1982). We suggest that ADL, MMSE and FAST assess different phenomena than pain, especially when patients with advanced dementia are included. We are not aware of pain treatment studies investigating the effect of analgesics on ADL functioning in NH patients with severe dementia. However, a recent review of physical inactivity and its relationship to pain (and vice versa) found a positive association between ADL and pain in older person without dementia (Plooij et al., 2012a). Another study on pain medication prescription and cognitive

function found no relationship between cognition (assessed by MMSE) and pain medication (Plooij et al., 2012b).

2.5 Intervention and control (related to Hypothesis 1)

The patients randomized to the intervention group followed a stepwise protocol for treatment of pain (SPTP) in accordance with recommendations of the American Geriatrics Society (American Geriatrics Society Panel on Pharmacological Management of Persistent Pain in Older Persons, 2009). Depending upon the ongoing medical treatment, patients in the intervention group received one step up of increased pain treatment, which might include acetaminophen oral (maximum increase to 3 g/day), extended release morphine oral (maximum 20 mg/day) or pregabalin oral (maximum 300 mg/day) using a fixed dose regime (Husebo et al., 2011b). Patients with swallowing difficulties were treated with buprenorphine transdermal plaster (maximum 10 µg/h for 7 days). Medication was offered by a registered nurse at breakfast, lunch and dinner (approximately 8:00 a.m., noon, 6:00 p.m.), respectively. If needed, combination therapy was allowed. Patients randomized to the control groups continued with treatment (including the ongoing pain treatment) and care as usual.

Related concomitant drugs, anti-dementia medication, psychotropic drugs, aspirin (one dose per day) or anti-inflammatory agent (e.g., ibuprofen) was allowed if remained stable for 4 weeks prior to study inclusion. Use of as-needed analgesics was allowed for all patients and was monitored during the study. Clinicians were advised when possible to keep prescriptions and doses of psychotropic medications unchanged. After the 12-week study period, the research team and a senior member of the NH staff from each institution reviewed the last assessment outcomes and drug prescriptions and optimized pain treatment for all patients.

2.6 Assessment tools (related to demographics and Hypotheses 2–6)

2.6.1 The MOBID-2 pain scale

Pain was assessed by the MOBID-2 pain scale at weeks 0, 2 and 4 following the MOBID-2 test procedures and guidelines described earlier (Husebo et al., 2010). In practice, the assessment of pain was based upon observation of the patient's immediate pain behaviour in connection with standardized, guided movements during morning care (MOBID-2, part 1, five items) and pain behaviour related to internal organs, head and skin (MOBID-2, part 2, five items) monitored over time. Primary caregivers, considered proxy raters, were encouraged to judge whether the patient's behaviour was related to pain or to dementia. For each item, the caregiver answered the question: 'How intense do you consider the pain to be?' and rated intensity on a numerical rating scale ranging from 0 (*no pain*) to 10 (*as bad as it possibly*

could be) (Jensen et al., 1994). MOBID-2 parts 1 and 2 and the total MOBID-2 scores are derived from the caregiver by overall intensity scores.

Test–retest reliability and inter- and intra-rater reliability of the MOBID-2 pain scale have been examined in previous studies, but different samples and time between assessments were applied (Husebo et al., 2007, 2009, 2010). In a study including 77 patients with severe dementia from one NH, the inter-rater and test–retest reliability (after 1 day) for pain intensity of items, subscales and total scale were mostly very good, with ICCs ranging from 0.80 to 0.94 and from 0.60 to 0.94, respectively (Husebo et al., 2010).

2.6.2 Cohen-Mansfield Agitation Inventory

The CMAI is a 29-item instrument (range 29–203 points) used by caregivers to rate the frequency with which NH patients with dementia manifest aggressive behaviour, physical non-aggressive behaviour or verbally agitated behaviour (Cohen-Mansfield and Libin, 2004). Behavioural disturbances were defined by a score of ≥ 39 . CMAI items are rated on a 1- to 7-point scale of frequency, ranging from never (1); occurring less than once a week (2); once or twice a week (3); several times a week (4); once or twice a day (5); several times a day (6); to several times an hour (7). Good reliability and validity has been reported (Cohen-Mansfield and Libin, 2004). The rating was made by a trained research assistant based upon a face-to-face interview with the caregiver who was familiar with the patient.

2.6.3 Neuropsychiatric Inventory – Nursing Home version

The NPI-NH is a 12-item instrument (range 0–144 points) developed to assist caregivers to rate the frequency and severity with which NH patients with dementia manifest neuropsychiatric symptoms such as delusions, hallucinations, agitation, depression, anxiety, euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, nocturnal behaviour, appetite and eating disorders (Selbaek et al., 2008). NPI-NH items are rated on a 1- to 4-point scale of frequency, ranging from (1) occasionally – less than once per week; (2) often – approximately once per week; (3) frequently – several times a week but less than every day; (4) very frequently – daily or essentially continuously present. The severity is rated as (1) mild – produces little distress in the patient; (2) moderate – more disturbing to the patient but can be redirected by the caregiver; (3) severe – very disturbing to the patient and difficult to redirect. Good reliability and validity of the Norwegian version of NPI-NH has been reported (Selbaek et al., 2008).

2.6.4 Barthel index of Activity of Daily Living

The Barthel ADL (range 0–20) evaluates the patient's physical function, which includes daily activities such as feeding,

moving, personal toilet and dressing (Mahoney and Barthel, 1965). The severity of functional impairment is rated as mild (12–20), moderate (9–10), severe (5–8) and very severe (<5). Reliability has been assessed thoroughly by self-report, by trained nurses and two independent skilled observers. Agreement was generally present in more than 90% of situations. Validity, reliability, sensitivity and clinical utility were excellent (Sheikh et al., 1979; Collin et al., 1988). Higher values indicate higher levels of activities of daily functioning and independence.

2.6.5 Mini-Mental State Examination

The MMSE is a 30-point scale that enables cut-off differentiation for levels of severity of cognitive impairment (i.e., 0–11 = severe, 12–17 = moderate, 18–23 = mild, 24–30 = no impairment) (Folstein et al., 1975). The questionnaire consists of several orientation question (10 points), registration and recall task (6), attention task (5), three-stage command (3), two naming tasks (2), repetition task (1), reading comprehension task (1), written sentence (1) and a visual construction (1). In this study, cognitive impairment was categorized as follows: moderate (MMSE score ≤ 19) and severe (MMSE score ≤ 12) (Burns et al., 2010).

Earlier reports suggested a cut-off point of 23/24 on the MMSE scale to be able to discriminate between patients with cognitive impairment and normal subjects (Folstein et al., 1975). However, latter studies found that the specificity of the MMSE was lower for individuals with less schooling and for those over the age of 65 (Mitchell, 2009). In addition, differences in application of the MMSE between primary care, memory clinic and community settings are obvious. To be sure to include patients with dementia in this study, we chose the latest recommendation by Burns et al., who categorized MMSE scores as having normal, mild, moderate (MMSE score ≤ 19) and severe (MMSE score ≤ 12) (Burns et al., 2010).

2.6.6 Functional Assessment Staging

FAST describes a continuum of seven successive stages and sub-stages from normality to most severe dementia (Hughes et al., 1982). Moderate to severe dementia is consistent with FAST stages of 5 or 6 or 7. Stage 5 is defined as moderately severe cognitive decline, with deficient performance in activities of daily living such as choosing proper clothing and maintaining hygiene. Stage 6 is defined as severe cognitive decline with incontinence and decreased ability to clothe, bathe, toilet oneself, severely limited speech, vocabulary and emotional expression. A patient who is considered to be at FAST stage 7 is no longer able to talk or smile, to walk and hold up her head. MMSE and FAST was used at weeks 0 and 8.

2.7 Analysis

Descriptive statistics was used to describe demographic characteristics and test scores of the study samples. Statistical

analyses were performed using the software program SPSS 17.0 (SPSS, Inc., Chicago, IL, USA).

Test–retest reliability was calculated using ICC_{2,1} between ratings at weeks 0 and 2, and between weeks 2 and 4. Each single MOBID-2 item of parts 1 and 2 and the total MOBID-2 pain scale score were analysed. ICC_{2,1} is a two-way mixed, absolute agreement model, also including variance due to systematic difference in the error variance (van der Roer et al., 2006). The SEM_{AGREEMENT} was estimated by taking the square root of the within-subject variance consisting of variance between the measures plus the residual variance. The SDC was calculated as SEM_{AGREEMENT} $\times \sqrt{2} \times 1.64$ (van der Roer et al., 2006; Kovacs et al., 2007), and is the test value that a patient must exceed to demonstrate an improvement above measurement error with 95% certainty (Bland and Altman, 1996).

Regarding responsiveness, Hypothesis 1 was examined by paired samples *t*-test. The 0-hypothesis of no difference in change between the two groups was examined by independent samples *t*-test, $p \leq 0.05$.

The hypothesis of moderate correlation ($0.50 > r \geq 0.30$; Pallant, 2005) between change scores of the total MOBID-2 pain scale versus CMAI (Hypothesis 2) and NPI-NH (Hypothesis 3) and weak or no correlation ($r < 0.30$) between total change scores of the MOBID-2 pain scale versus measures of ADL, MMSE and FAST (Hypotheses 4–6) were examined by Pearson's correlation.

3. Results

3.1 Demographics

At baseline, 163 patients who were not expected to change (control group) were included in the test–retest reliability study, and 203 patients with MOBID-2 pain scale scores ≥ 3 participated in the responsiveness study; 99 patients from the control group and 104 from the intervention group.

The demographic and test characteristics of the patients at baseline are shown in Table 1. These characteristics were rather similar in patients included in the reliability and responsiveness studies; most were women, 78% and 74%, with a mean age of 87 and 85 years, respectively. During the 8 weeks, 20 and 28 patients were lost in the control and the SPTP groups, respectively ($p = 0.298$) (Husebo et al., 2011b). There were 14 deaths during the study period: 8 in the control and 6 the intervention group.

3.2 Clinical pain scores in NH patients with dementia

As expected, the mean MOBID-2 pain scale total scores were higher at baseline in the responsiveness study (mean score 5.4) than in the reliability study

Table 1 Demographic and clinical characteristics of patients who completed MOBID-2 pain score at baseline and participated in the study of test–retest reliability (control patients) and responsiveness (patients with MOBID-2 scores ≥ 3).

	Rest–retest reliability (<i>n</i> = 163)	Responsiveness (<i>n</i> = 203)
Female, <i>n</i> (%)	128 (78)	149 (74)
Age (years), mean (SD)	86.5 (6.7)	85.4 (6.9)
MOBID-2 pain scale, ^a mean (SD)	3.7 (2.4)	5.4 (1.8)
CMAI, ^b mean (SD)	56.1 (16.0)	56.2 (14.5)
NPI-NH, ^c mean (SD)	31.4 (21.4)	34.5 (21.7)
FAST, ^d mean (SD)	6.0 (0.7)	6.0 (0.7)
MMSE, ^e mean (SD)	8.4 (6.7)	8.1 (6.7)
ADL, ^f mean (SD)	8.6 (5.6)	7.6 (5.6)
ICD diagnoses	5.4 (2.1)	5.6 (2.0)

^aMobilization-Observation-Behavior-Intensity-Dementia-2 pain scale (range 0–10); higher scores mean more pain.

^bCohen-Mansfield Agitation Inventory (range 29–203); higher scores mean more agitation (scores ≥ 39 usually accepted as clinically significant).

^cNeuropsychiatric Inventory – Nursing Home version (range 1–144); higher scores mean more psychiatric symptoms.

^dFunctional Assessment Staging (range 1–7); higher scores mean more cognitive impairment.

^eMini-Mental State Examination (range 0–30); higher scores mean more cognitive impairment.

^fActivities of Daily Living (range 0–20); higher scores mean higher physical function.

(mean score 3.7) (Table 2). Guided movements of the hips, knees and ankles and turning over in bed were the most painful items related to the musculoskeletal system, MOBID-2 part 1. The most painful locations that might be related to internal organs, head and skin (MOBID-2 part 2) were the areas for the pelvis and genital organs. Table 2 shows the scores of single MOBID-2 pain scale items and total score.

3.3 Reliability and measurement error

Test–retest reliability between baseline and 2 weeks was high for the separate items ($ICC_{2,1} = 0.731–0.857$) and the total score ($ICC_{2,1} = 0.805$). Between weeks 2 and 4, reliability was even better for the separate items ($ICC_{2,1} = 0.729–0.889$) and for the total score ($ICC_{2,1} = 0.852$) (Table 3). The SEM and the SDC for

the MOBID-2 total score improved from 1.9 and 3.1 (0–2 weeks) to 1.4 and 2.3 (2–4 weeks) over time.

3.4 Responsiveness

Supporting our first hypothesis, patients randomized to stepwise pain treatment improved more on MOBID-2 pain scale for single items, MOBID-2 parts 1 and 2, and on the total score than did the control group that received treatment and care as usual. Test scores for the two groups are shown in Table 4 and demonstrate a mean improvement on MOBID-2 total score of 1.7 in the intervention group and 0.3 in the control group. The difference in change was found to be statistically significant, in favour of the intervention group for MOBID-2 part 1 ($p < 0.001$), MOBID-2 part 2 ($p < 0.001$) and MOBID-2 total score ($p < 0.001$).

Table 2 Pain scores on MOBID-2 single items and total scores of patients who participated in the reliability study (controls with no intervention) and in the responsiveness study (MOBID-2 ≥ 3).

	Reliability (<i>n</i> = 163) MOBID-2, mean (SD)	Responsiveness (<i>n</i> = 203) MOBID-2 ≥ 3 , mean (SD)
1. Hands	0.8 (1.9)	1.4 (2.3)
2. Arms	1.7 (2.3)	2.4 (2.7)
3. Legs	2.6 (2.9)	3.3 (3.0)
4. Turn over	1.9 (2.5)	2.9 (2.9)
5. Sit	1.6 (2.2)	2.6 (2.8)
6. Head, mouth, neck	1.0 (1.8)	1.7 (2.3)
7. Heart, lung, chest wall	0.8 (1.7)	1.1 (1.9)
8. Abdomen	0.9 (1.8)	1.3 (2.1)
9. Pelvis, genital organ	1.6 (2.5)	2.4 (2.9)
10. Skin	1.7 (2.3)	2.3 (2.7)
11. Overall pain score	3.7 (2.4)	5.4 (1.8)

Table 3 Test–retest reliability of MOBID-2 pain scale data between baseline (week 0) and week 2, and weeks 2 and 4 by intraclass correlation coefficient (ICC_{2,1}), standard error of measurement (SEM) and smallest detectable change (SDC), assessed in controls who received treatment as usual.

	Weeks 0–2 (<i>n</i> = 163)			Weeks 2–4 (<i>n</i> = 159)		
	ICC _{2,1}	SEM	SDC	ICC 2.1	SEM	SDC
1. Hand	0.852	1.1	1.8	0.889	0.9	1.4
2. Arms	0.731	1.6	2.6	0.785	1.4	2.3
3. Legs	0.820	1.6	2.6	0.879	1.2	1.9
4. Turn over	0.841	1.3	2.1	0.827	1.3	2.1
5. Sit	0.813	1.2	1.9	0.828	1.2	1.9
6. Head, mouth, neck	0.760	1.1	1.8	0.729	1.1	1.8
7. Heart, lung, chest wall	0.811	0.9	1.4	0.861	0.7	1.1
8. Abdomen	0.757	1.1	1.8	0.815	0.9	1.4
9. Pelvis, genital organs	0.732	1.6	2.6	0.855	1.2	1.9
10. Skin	0.857	1.2	1.9	0.822	1.3	2.1
MOBID-2 total score	0.805	1.9	3.1	0.852	1.4	2.3

Our second hypothesis was confirmed as we found a moderate association between change scores of MOBID-2 pain scale and change scores of CMAI (*n* = 90) in the intervention group after 8 weeks. Our third hypothesis was not confirmed as we found only a weak association between change scores of MOBID-2 and NPI-NH (Table 5). Hypotheses 4–6 were confirmed as there was no correlation between change scores on MOBID-2 pain scale and change scores on ADL, MMSE and FAST in the intervention group (Table 5).

4. Discussion

In this study, we assessed the following measurement properties of the MOBID-2 pain scale in NH patients with advanced dementia: test–retest reliability, SEM and responsiveness to change. The study was planned and data were collected alongside a large cluster randomized clinical trial (Husebo et al., 2011b). Reliability values by ICCs were found to be high for both time intervals of 2 weeks. As the patients randomized to control received treatment and care as usual, they were not expected to change systematically. Measurement error by SEM and SDC values ranged between 1.4 and 1.9 and 3.1 and 2.3, respectively. This indicates that a decrease of at least 3 on the 0–10 point total MOBID-2 pain scale is needed to be confident that an improvement in individual patients is not merely a measurement error.

Further, we found evidence that the total MOBID-2 pain scale is responsive to a decrease in pain. The hypothesis that patients with pain who received step-

Table 4 Test scores of single MOBID-2 items, MOBID-2 parts 1 and 2, and MOBID-2 total score at baseline and at week 8 for patients with pain at baseline (MOBID-2 ≥ 3), and difference in score changes in controls receiving treatment as usual and in patients receiving increased pain treatment (intervention).

Item	Control (<i>n</i> = 99)			Intervention (<i>n</i> = 104)			Difference in change <i>p</i> -value ^b
	Baseline, mean (SD)	Week 8, mean (SD)	Change, ^a mean (SD)	Baseline, mean (SD)	Week 8, mean (SD)	Change, ^a mean (SD)	
1. Hands	1.2 (2.3)	1.2 (2.3)	0.0 (1.2)	1.5 (2.3)	1.0 (1.8)	0.4 (2.1)	0.577
2. Arms	2.5 (2.7)	2.4 (2.8)	0.1 (2.6)	2.5 (2.8)	1.3 (2.0)	1.1 (2.4)	0.005
3. Legs	3.7 (3.1)	2.9 (2.9)	0.9 (2.8)	2.9 (2.9)	2.0 (2.2)	0.8 (3.0)	0.032
4. Turn over	2.8 (2.8)	2.6 (2.7)	0.3 (2.6)	3.0 (2.9)	1.6 (2.0)	1.3 (2.7)	0.013
5. Sit	2.2 (2.3)	2.3 (2.5)	-0.2 (2.1)	3.1 (3.1)	1.6 (2.0)	1.5 (1.3)	0.027
6. Head, mouth, neck	1.3 (2.1)	1.0 (2.0)	0.3 (1.6)	1.8 (2.4)	0.8 (1.6)	1.0 (2.5)	0.523
7. Heart, lung, chest wall	1.1 (2.0)	1.1 (1.8)	0.0 (1.7)	1.1 (1.9)	0.5 (1.2)	0.6 (1.7)	0.011
8. Abdomen	1.1 (2.1)	0.9 (1.8)	0.1 (0.2)	1.4 (2.1)	0.4 (1.0)	1.0 (2.3)	0.029
9. Pelvis, genital organ	2.0 (2.8)	2.0 (2.6)	0.1 (3.0)	2.5 (2.9)	0.9 (1.6)	1.6 (2.5)	0.003
10. Skin	2.3 (2.6)	1.9 (2.7)	2.2 (2.8)	1.1 (2.8)	1.1 (1.9)	1.0 (3.1)	0.047
MOBID-2 part 1 (items 1–5)	5.0 (2.7)	4.5 (2.9)	0.5 (2.8)	4.8 (2.8)	2.8 (2.4)	2.0 (3.2)	<0.001
MOBID-2 part 2 (items 6–10)	4.2 (2.7)	3.8 (2.9)	0.4 (2.8)	4.6 (2.5)	2.4 (2.2)	2.2 (3.1)	<0.001
MOBID-2 total pain score	5.3 (1.8)	4.3 (2.5)	1.0 (2.2)	5.4 (2.0)	2.6 (2.1)	2.7 (2.7)	<0.001

^aStudent's paired sample *t*-test; change from baseline to week 8.

^bIndependent-samples *t*-test; difference in change between groups.

Table 5 Testing hypothesis of correlation (Pearson, r) between change scores of MOBID-2 and change scores of CMAI, NPI-NH, FAST, MMSE and ADL after 8 weeks of pain treatment, $n = 104$.

Measurement	MOBID-2	Expected range in correlation	Hypothesis confirmed?
CMAI	0.351*	$0.60 > r \geq 0.30$	Yes
NPI-NH	0.275*	$0.60 > r \geq 0.30$	No
FAST	-0.030	$r < 0.20$	Yes
MMSE	-0.089	$r < 0.20$	Yes
ADL	-0.081	$r < 0.20$	Yes

ADL, Activities of Daily Living; CMAI, Cohen-Mansfield Agitation Inventory; FAST, Functional Assessment Staging; MMSE, Mini-Mental State Examination; MOBID-2, Mobilization-Observation-Behavior-Intensity-Dementia-2 pain scale; NPI-NH, Neuropsychiatric Inventory – Nursing Home version.

* $p < 0.001$.

wise increased pain medication over time obtained more pain relief than did patients who received treatment and care as usual was confirmed. This difference in change between the patient groups also applied to MOBID-2 parts 1 and 2, and to most of the single MOBID-2 items. Other hypotheses exploring the responsiveness (longitudinal validity) provided evidence that a decrease in total MOBID-2 pain scale scores reflects a decrease in aspects related to pain. Results support the instrument's ability to act as an outcome measure of pain in efficacy studies of pain treatment including patients with advanced dementia. This is of key importance because a central step to ensure improved pain management in patients with dementia is the ability of a pain tool to capture the effect of pain treatment over time.

The association between MOBID-2 pain scale and CMAI was, as expected, moderately and confirmed our hypothesis. Pain behaviour certainly includes elements of behavioural disturbances in patients with dementia, and earlier findings demonstrated reductions in agitation and aggression by systematic pain management in patients with dementia (Husebo et al., 2011b). Especially, verbal agitation behaviours, such as complaining or negativism, and physical non-aggressive behaviour, such as pacing and restlessness, responded to this treatment (Husebo et al., 2013).

Testing the hypothesis of correlation between change scores of MOBID-2 and scores of NPI-NH turned out to be lower than expected and demonstrated only weak correlation. Thus, a change in pain scores seems to be a different phenomenon than changes in the overall score of NPI-NH, which, besides agitation and mood syndrome, also includes psychiatric symptoms such as delusions, hallucinations or disinhibition.

For MMSE and ADL, we found, as expected, that changes in pain is truly a different phenomenon than changes in scores on ADL functioning or cognition. It has been suggested earlier that patients exhibit higher

levels of general activity during treatment with acetaminophen (Chibnall et al., 2005). However, the sample size of that study was quite small ($n = 25$) and the ADL index was used as an outcome measure. More detailed analyses are needed to investigate the impact of different analgesic groups on ADL functioning and cognition in patients with advanced dementia.

4.1 Pain frequency and location

Our data confirm that pain is still a frequent symptom in NH patients with dementia. Sixty-two percent of the sample was found to be in pain defined as MOBID-2 pain scale scores ≥ 3 . A previous cross-sectional study from a NH with a high focus upon palliative care, with physicians and caregivers skilled in pain treatment, reported a pain prevalence of 54% (Husebo et al., 2007). Supporting previous results, most frequent and painful were mobilizing the hips and legs (Husebo et al., 2010). As the most common clinical syndromes involving somatic nociceptor activities are associated with degenerative diseases or other painful conditions in muscles, bone and/or joints, gently guided movements, for instance, performed during morning care, should be part of a routine assessment protocol to reveal musculoskeletal pain in these individuals (Husebo et al., 2010; Corbett et al., 2012). Pain from internal organs, head and skin was less frequently observed; most frequent were painful conditions in the pelvis and/or genital organs and pain that might originate from the skin. In general, pain seemed attributed mainly to the musculoskeletal system. In general, diagnosing the type of pain, i.e., nociceptive, peripheral and central neuropathic syndrome, remains extremely difficult, especially in patients in an advanced stage and it was recently suggested that central neuropathic pain is by far the most under-treated type of pain in patients with dementia (Scherder and Plooi, 2012).

4.2 Reliability and measurement error

Sufficient test–retest reliability is a prerequisite for responsiveness to change of a measure (Terwee et al., 2007; Strand et al., 2011). High to excellent inter- and intra-tester and test–retest reliability of the MOBID-2 has been demonstrated in previous studies based upon scores from both video recordings and bedside care situations in patients with severe dementia (Husebo et al., 2009, 2010). Compared with these studies, test–retest reliability values of the present study were somewhat less favourable, probably due to the longer time interval applied. Although the patients received treatment and care as usual, their conditions might still fluctuate over time, and pain medication or other treatments might be initiated or terminated during the time periods of 2 weeks between the assessments. The most important information derived from this study is the SDC, finding that a decrease of at least 3 on the 0–10 point total MOBID-2 pain scale is needed to demonstrate a change above measurement error in individual patients. Taking the two test periods into consideration, we find the stability of reliability estimates to be satisfactory.

It is noteworthy that the test–retest reliability of the item for head, mouth and neck was low. In addition, this item did not seem responsive to change after individual pain treatment. Orofacial pain is a complex clinical symptom of temporomandibular disorders characterized by a reduction of chewing ability. In turn, impaired chewing may result in chronic malnutrition, vitamin deficiency and poorer physical activities (Lobbezoo et al., 2011). A recent review article presents the complexity of this vicious circle and underlines the needs for the development and testing of a new tool that can help dentists in the diagnosis of orofacial and dental pain in this vulnerable patient population (Lobbezoo et al., 2011). It is questionable whether the MOBID-2 pain scale can be further developed to cover all different aspects of pain, such as orofacial pain and neuropathic pain in patients with diabetes or after stroke. The European COST-Action TD 1005, *Assessment of Pain in Patients with Cognitive Impairment, especially Dementia*, is presently working on the development of a comprehensive and internationally agreed-on assessment toolkit for older adults targeting the various aspects of pain and different pain diagnoses.

4.3 Responsiveness

Three other studies had previously investigated the responsiveness of pain tools for patients with demen-

tia or non-verbally communicating elderly. Cohen-Mansfield and Lipson (2008) compared the responsiveness of six inform rating and observation pain tools in order to assess pain treatment effect in 36 patients with dementia. Another subsequent trial of pain treatment in non-verbally communicating elderly by Morello et al. (2007) reported very good responsiveness of the Elderly Pain Caring Assessment (EPCA-2) after pain intervention in 283 non-verbally communicating elder patients. The third study by Rat et al. (2011) included 91 acute pain patients with an inability to communicate verbally, to investigate the responsiveness of the Algoplus®.

Although these studies report promising results, the methodological approaches are questionable in light of the latter requirements by the COSMIN group, which recommends a checklist for assessing the methodological quality of health status measurement instruments (Mokkink et al., 2010). Following this, none of these studies examined measurement error by the SEM and the SDC, important parameters for judging change scores. In studies without an available gold standard, responsiveness measures cannot be based upon a ROC analysis, but rather on hypotheses about changes in scores formulated a priori. Some of the studies were underpowered, lack a sample sizes calculation or drop-out rate, and/or control groups to compare changes over time with the intervention groups. Individual pain treatment is necessary to obtain relevant change scores, but interventions are often not described precisely.

Until recently, there was no consensus in the research literature on what constitutes a responsive measure and how responsiveness should be examined and quantified. Based upon a Delphi process, the COSMIN group concluded that responsiveness relies upon differences between changes in scores, or expected correlations between changes in scores on the respective instrument, and other instruments known to have adequate responsiveness (Mokkink et al., 2010). In the current study, we followed the recommendations of the American Geriatrics Society for pharmacological management of persistent pain in older persons and presumed pain medication as a gold standard for change in pain (American Geriatrics Society Panel on Pharmacological Management of Persistent Pain in Older Persons, 2009). Thus, when testing our first hypothesis, we expected more decrease in pain in patients with pain who received increased pain medication, compared to those who continued with treatment and care as usual. This hypothesis was confirmed as the intervention group obtained a mean pain decrease of 2.7 on the 10 point

total MOBID-2 pain scale, while it decreased by only a mean of 1.0 in the control group.

4.4 Ethical considerations

It may be an ethical concern that patients randomized to control were not treated at once in according with a stepwise protocol for treating pain by the research team. The risk associated with individual or cluster randomization is particularly great when the intervention group receives important treatment such as pain management. The challenge of designing an ethical randomized trial requires balancing the potential benefits and risk of harm faced by individual participation with the potential long-term benefit to those subjects and to society at large (Donner and Klar, 2000). Thereby, clinical trials should meet the criteria of minimal risk, which means that the probability and magnitude of harm or discomfort anticipated are not greater than those ordinarily encountered in daily life or during the performance of routine physical examinations or tests (US Federal Government – Office for Protection from Research Risks, 1994).

Concerning the risk/benefit ratio of the control group in the present study, we attempted to ensure that participants benefitted by participation in this study. All patients received optimized treatment by an expert group based upon data collected during enrolment and before randomization, as well as after the end of the study. This means that the patients were taken better care of regarding pain treatment than if they had not participated in the study. Use of as-need analgesics was allowed for all patients and none were removed from their ongoing pain treatment. NH staff improved their knowledge regarding pain assessment and –treatment by standardized education.

4.5 Limitations of the study

The current RCT study is the first adequately powered parallel group RCT of pain management for patients with advanced dementia, providing a good basis for this responsiveness study, and following the latest COSMIN recommendations as to how measurement properties should be examined. However, our study has several limitations. The COSMIN checklist has not yet been reliability tested and some of the standards need further refinement, e.g., by defining an adequate sample size or test–retest time interval or when something is adequately described (Mokkink et al., 2010). In this study, pain improved in both intervention and control clusters (Husebo et al., 2011a). That may indicate a Hawthorne effect perhaps related to increased

staff competence and training. However, use of as-need analgesics was allowed for all patients and should have been monitored more closely during the study. Precautions were taken to blind research assistants and primary caregivers for group allocation, but despite these efforts, fully blind these studies will always be difficult because of the whole NH setting.

Although enormous under-treatment of pain is still a challenge in NH patients with dementia (Husebo et al., 2008; Plooij et al., 2012b), these individuals are at high risk for poly-pharmacy and related side effects. Recent literature indicate that analgesic use is almost higher among people with dementia compared with older adults without dementia, although the majority of these studies were undertaken in Scandinavia and may not be indicative of treatment elsewhere (Lovheim et al., 2008; Haasum et al., 2011). The frequency of analgesic prescription does not necessarily indicate whether an appropriate analgesic treatment is being prescribed to the right people at the right time (Corbett et al., 2012). Among individuals with dementia, this will depend, to a great extent, upon the timely identification and assessment of pain and pain treatment effect. In these patients, we must find the balance between effect and side effect, as they are unable to report neither pain nor the effect of pain management or side effects of poly-pharmacy.

Clinical study registration

The trial is registered at ClinicalTrials.gov, number NCT01021696 and at the Norwegian Medicines Agency (EudraCTnr: 2008–007490-20).

Ethical approval

Regional Committee for Medical Research Ethics, Western Norway (REK-Vest nr: 248.08).

Author contributions

B.S.H., R.O. and L.I.S. conceived the study design. B.S.H. was the primary investigator of the study. All authors contributed to the statistical analysis, interpretation of the data and discussion of the results, and they drafted, commented and wrote the manuscript.

Acknowledgements

We thank the patients, their relatives and the NH staff for their willingness and motivation, which made this study possible. B.S.H. and L.I.S. acknowledge support from the COST programme (European Cooperation in the Field of

Scientific and Technical Research) for COST-Action TD 1005, Assessment of Pain in Patients with Mental Impairment, especially Dementia.

References

- Arola, H.M., Nicholls, E., Mallen, C., Thomas, E. (2010). Self-reported pain interference and symptoms of anxiety and depression in community-dwelling older adults: Can a temporal relationship be determined? *Eur J Pain* 14, 966–971.
- Ayalon, L. (2009). Willingness to participate in Alzheimer disease research and attitudes towards proxy-informed consent: Results from the health and retirement study. *Am J Geriatr Psychiatry* 17, 65–74.
- Bland, J.M., Altman, D.G. (1996). Measurement error. *BMJ* 312, 1654.
- Burns, A., Karim, S., Morris, J., Byrne, J. (2010). The conversational mini-mental state examination: Estimating cognitive impairment. *Aging Ment Health* 14, 692–694.
- Chen, J., Devine, A., Dick, I.M., Dhaliwal, S.S., Prince, R.L. (2003). Prevalence of lower extremity pain and its association with functionality and quality of life in elderly women in Australia. *J Rheumatol* 30, 2689–2693.
- Chibnall, J.T., Tait, R.C., Harman, B., Luebbert, R.A. (2005). Effect of acetaminophen on behavior, well-being, and psychotropic medication use in nursing home residents with moderate-to-severe dementia. *J Am Geriatr Soc* 53, 1921–1929.
- Cohen-Mansfield, J., Libin, A. (2004). Assessment of agitation in elderly patients with dementia: Correlations between informant rating and direct observation. *Int J Geriatr Psychiatry* 19, 881–891.
- Cohen-Mansfield, J., Lipson, S. (2008). The utility of pain assessment for analgesic use in persons with dementia. *Pain* 134, 16–23.
- Cohen-Mansfield, J., Taylor, L. (1998). The relationship between depressed affect, pain and cognitive function: A cross-sectional analysis of two elderly populations. *Aging Ment Health* 2, 313–318.
- Collin, C., Wade, D.T., Davies, S., Horne, V. (1988). The Barthel ADL index: A reliability study. *Int Disabil Stud* 10, 61–63.
- Corbett, A., Husebo, B.S., Malcangio, M., Staniland, A., Cohen-Mansfield, J., Aarsland, D., Ballard, C. (2012). Assessment and treatment of pain in people with dementia. *Nat Rev Neurol* 8, 264–274.
- Donner, A., Klar, N. (2000). *Design and Analysis of Cluster Randomized Trials in Health Research* (Chichester, West Sussex: John Wiley & Sons Ltd).
- Folstein, M.F., Folstein, S.E., McHugh, P.R. (1975). 'Mini-mental state'. A practical method for grading cognitive state of patients for clinician. *J Psychiatr Res* 12, 189–198.
- Haasum, Y., Fastbom, J., Fratiglioni, L., Kareholt, I., Johnell, K. (2011). Pain treatment in elderly persons with and without dementia a population-based study of institutionalized and home-dwelling elderly. *Drugs Aging* 28, 283–293.
- Hadjistavropoulos, T., Herr, K., Turk, D.C., Fine, P.G., Dworkin, R.H., Helme, R., Jackson, K., Parmelee, P., Rudy, T., Lynn, B., Chibnall, J., Craig, K., Ferrell, B., Ferrell, B., Fillingim, R., Gagliese, L., Gallagher, R., Gibson, S., Harrison, E., Katz, B., Keefe, F., Lieber, S., Lussier, D., Schmader, K., Tait, R., Weiner, D., Williams, J. (2007). An interdisciplinary expert consensus statement on assessment of pain in older persons. *Clin J Pain* 23, S1–S43.
- Herr, K. (2011). Pain assessment strategies in older patients. *J Pain* 12, S3–S13.
- Hughes, C.P., Berg, L., Danziger, W.L., Coben, L.A., Martin, R.L. (1982). A new clinical scale for the staging of dementia. *Br J Psychiatry* 140, 566–572.
- Husebo, B.S., Ballard, C., Aarsland, D. (2011a). Pain treatment of agitation in patients with dementia: A systematic review. *Int J Geriatr Psychiatry* 26, 1012–1018.
- Husebo, B.S., Ballard, C., Cohen-Mansfield, J., Nilsen, O.B., Aarsland, D. (2013). The response of agitated behavior to pain management in persons with dementia. *Am J Geriatr Psychiatry*. doi: 10.1016/j.jagp.2012.12.006; [Epub ahead of print].
- Husebo, B.S., Ballard, C., Sandvik, R., Nilsen, O.B., Aarsland, D. (2011b). Efficacy of treating pain to reduce behavioural disturbances in residents of nursing homes with dementia: Cluster randomised clinical trial. *BMJ* 343, 1–10.
- Husebo, B.S., Kunz, M., Achterberg, W.P., Lobbezoo, F., Kappesser, J., Tudose, C., Lautenbacher, S. (2012). Pain assessment and treatment challenges in patients with dementia. *Zeitschr Neuropsychol* 23, 237–244.
- Husebo, B.S., Strand, L.I., Moe-Nilssen, R., Husebo, S.B., Aarsland, D., Ljunggren, A.E. (2008). Who suffers most? Dementia and pain in nursing home patients: A cross-sectional study. *J Am Med Dir Assoc* 9, 427–433.
- Husebo, B.S., Strand, L.I., Moe-Nilssen, R., Husebo, S.B., Ljunggren, A.E. (2009). Pain behaviour and pain intensity in older persons with severe dementia: Reliability of the MOBID pain scale by video uptake. *Scand J Caring Sci* 24, 380–391.
- Husebo, B.S., Strand, L.I., Moe-Nilssen, R., Husebo, S.B., Ljunggren, A.E. (2010). Pain in older persons with severe dementia. Psychometric properties of the mobilization-observation-behaviour-intensity-dementia (MOBID-2) pain scale in a clinical setting. *Scand J Caring Sci* 24, 380–391.
- Husebo, B.S., Strand, L.I., Moe-Nilssen, R., Husebo, S.B., Snow, A.L., Ljunggren, A.E. (2007). Mobilization-observation-behavior-intensity-dementia pain scale (MOBID): Development and validation of a nurse-administered pain assessment tool for use in dementia. *J Pain Symptom Manage* 34, 67–80.
- Jakobsson, U., Hallberg, I.R. (2002). Pain and quality of life among older people with rheumatoid arthritis and/or osteoarthritis: A literature review. *J Clin Nurs* 11, 430–443.
- Jensen, M.P., Turner, J.A., Romano, J.M. (1994). What is the maximum number of levels needed in pain intensity measurement? *Pain* 58, 387–392.
- Kovach, C.R., Logan, B.R., Noonan, P.E., Schlidt, A.M., Smerz, J., Simpson, M., Wells, T. (2006). Effects of the serial trial intervention on discomfort and behavior of nursing home residents with dementia. *Am J Alzheimers Dis Other Demen* 21, 147–155.
- Kovacs, F.M., Abaira, V., Royuela, A., Corcoll, J., Alegre, L., Cano, A. (2007). Minimal clinically important change for pain intensity and disability in patients with nonspecific low back pain. *Spine* 32, 2915–2920.
- Lobbezoo, F., Weijnen, R.A.F., Scherder, E.J.A. (2011). Topical review: Orofacial pain in dementia patients. A diagnostic challenge. *J Orofac Pain* 25, 6–14.
- Lovheim, H., Karlsson, S., Gustafson, Y. (2008). The use of central nervous system drugs and analgesics among very old people with and without dementia. *Pharmacoepidemiol Drug Safety* 17, 912–918.
- Mahoney, F.I., Barthel, D.W. (1965). Functional evaluation: The Barthel index. *Md State Med J* 14, 61–65.
- Manfredi, P.L., Breuer, B., Wallenstein, S., Stegmann, M., Bottomley, G., Libow, L. (2003). Opioid treatment for agitation in patients with advanced dementia. *Int J Geriatr Psychiatry* 18, 700–705.
- Mitchell, A.J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *J Psychiatry Res* 43, 411–431.
- Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., Bouter, L.M., de Vet, H.C.W. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res* 19, 539–549.
- Morello, R., Jean, A., Alix, M., Sellin-Peres, D., Fermanian, J. (2007). A scale to measure pain in non-verbally communicating older patients: The EPCA-2 study of its psychometric properties. *Pain* 138, 87–98.
- Office for Protection from Research Risks (1994). *Protection of Human Subjects, Title 45, Code of Federal Regulations, Part 46* (Bethesda, MD: Department of health and Human Services NIH).
- Pallant, J. (2005). *SPSS Survival Manual. Correlation (Chapter 11)*, 2nd edn (Berkshire: Open University Press).
- American Geriatrics Society Panel on Pharmacological Management of Persistent Pain in Older Persons (2009). Pharmacological management of persistent pain in older persons. *J Am Geriatr Soc* 57, 1331–1346.
- Plooijs, B., Scherder, E.J.A., Eggermont, L.H.P. (2012a). Physical inactivity in aging and dementia: A review of its relationship to pain. *J Clin Nurs* 21, 3002–3008.

- Plooij, B., van der Spek, K., Scherder, E.J.A. (2012b). Pain medication and global cognitive functioning in dementia patients with painful conditions. *Drugs Aging* 29, 377–384.
- Rat, P., Jouve, E., Pickering, G., Donnarel, L., Nguyen, L., Michel, M., Capriz-Ribièrel, F., Lefebvre-Chapiro, S., Gauquelin, F., Bonin-Guillaumel, S. (2011). Validation of an acute pain-behavior scale for older persons with inability to communicate verbally: Algoplus. *Eur J Pain* 15, 1–10.
- van der Roer, N., Ostelo, R.W.J.G., Bekkering, G.E., Van Tulder, M.W., de Vet, H.C.W. (2006). Minimal clinically important change for pain intensity, functional status, and general health status in patients with non-specific low back pain. *Spine* 31, 578–582.
- Scherder, E.J.A., Plooij, B. (2012). Assessment and management of pain, with particular emphasis on central neuropathic pain, in moderate to severe dementia. *Drugs Aging* 29, 701–706.
- Selbaek, G., Kirkevold, O., Sommer, O.H., Engedal, K. (2008). The reliability and validity of the Norwegian version of the neuropsychiatric inventory, nursing home version (NPI-NH). *Int Psychogeriatr* 20, 375–382.
- Sheikh, K., Smith, D.S., Meade, T.W., Goldenberg, E., Brennan, P.J., Kinsella, G. (1979). Repeatability and validity of a modified activities of daily living (ADL) index in studies of chronic disability. *Int Rehabil Med* 1, 51–58.
- Strand, L.I., Anderson, B., Lygren, H., Skouen, J.S., Ostelo, R., Magnussen, L.H. (2011). Responsiveness to change of 10 physical tests used for patients with back pain. *Phys Therapy* 91, 404–415.
- Terwee, C.B., Bot, S.D.M., De Boer, M.R., van der Windt, D.A.W.M., Knol, D.L., Dekker, J., Bouter, L.M., de Vet, H.C.W. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60, 34–42.