Act in case of depression!

Validation and effectiveness of a multidisciplinary depression care program in nursing homes

ISBN: 978-94-6203-052-7

Act in case of Depression! Validation and effectiveness of a multidisciplinary depression care program in nursing homes

Thesis with summary in Dutch, *Radboud University Nijmegen*

Nijmegen, 2012

Printed and bound by CPI Wöhrmann print service

Proofreading by K. Scott Dunn

Lay out and design cover by Ruslan Leontjevas

This thesis was prepared within the department of *Primary and Community Care, Centre for Family Medicine, Geriatric Care and Public health*, Radboud University Nijmegen, Medical Centre, the Netherlands. This department is part of the Nijmegen Centre for Evidence Based Practice (NCEBP), one of the research institutes of the Radboud University Nijmegen.

The following organisations of the Nijmegen University Network of Nursing Homes (Universitair Kennisnetwerk Ouderenzorg Nijmegen, UKON, www.uko-n.nl) participated in the studies: Archipel (www.archipelzorggroep.nl), Kalorama (www.kalorama.nl), De Riethorst Stromenland (www.deriethorststromenland.nl), SVRZ (www.svrz.nl), Volckaert-SBO (www.volckaert-sbo.nl), Vitalis Woonzorg groep (www.vitalisgroep.nl), de Waalboog (www.waalboog.nl), De Wever (www.dewever.nl), de Zorgboog (www.zorgboog.nl), De Zorggroep (www.dezorggroep.nl),

The studies described in this thesis were financed by the Netherlands Organisation for Health Research and Development (ZonMw, grant number 170992801). Financial support for the printing of this thesis has been kindly provided by the department of *Primary and Community Care, Centre for Family Medicine, Geriatric Care and Public health*, Alzheimer Nederland (Bunnik), and Nederlandse Vereniging voor Gerontologie.

ZonMw

UMC St Radboud

alzheimer nederland

NVG

Act in case of Depression!

Validation and effectiveness of a multidisciplinary depression care program in nursing homes

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. mr. S.C.J.J. Kortmann,
volgens besluit van het college van decanen
in het openbaar te verdedigen op maandag 24 september 2012
om 13.00 uur precies

door

Ruslanas Leontjevas

Promotoren:	prof. dr. R.T.C.M. Koopmans
	prof. dr. M.J.F.J. Vernooij-Dassen
Copromotoren:	dr. D.L. Gerritsen
	dr. M. Smalbrugge (VU Medisch Centrum, Amsterdam)
Manuscriptcommissie:	prof. dr. A.W.M. Evers
	prof. dr. A.E.M. Speckens
	prof. dr. A.L. Francke (VU Medisch Centrum, Amsterdam)

Act in case of Depression!

Validation and effectiveness of a multidisciplinary depression care program in nursing homes

Doctoral Thesis

To obtain the degree of doctor

From Radboud University Nijmegen

On the authority of the Rector Magnificus

prof. dr. S.C.J.J. Kortmann,

According to the decision of the Council of Deans

To be defended in public on Monday, September 24, 2012

at 13.00 hours

by

Ruslanas Leontjevas

Born on March 4, 1972

in Klaipėda, Lithuania

Supervisors: prof. dr. R.T.C.M. Koopmans

prof. dr. M.J.F.J. Vernooij-Dassen

Co-supervisors: dr. D.L. Gerritsen

dr. M. Smalbrugge (VU University Medical Center, Amsterdam)

Doctoral Thesis Committee:

prof. dr. A.W.M. Evers

prof. dr. A.E.M. Speckens

prof. dr. A.L. Francke (VU University Medical Center, Amsterdam)

Contents

Introduction section

Chapter

1

General Introduction

Introduction

This thesis focuses on the improvement of depression care in nursing homes. In order to improve the detection and treatment of depression and depressive symptoms in nursing home (NH) residents, NH professionals of the Nijmegen University Network of NHs (UKON*) developed a multidisciplinary evidence and practice-based care program Act in case of Depression (AiD, in Dutch: *Doen bij Depressie*). A two-year multicenter randomized controlled trial was conducted to provide insight into the effectiveness of the care program. In this thesis, results of the evaluation of the validity of the intervention and its effects are presented.

The thesis consists of an *Introduction Section,* including this chapter and the initial protocol of the AiD trial (Chapter 2); a *Results Section,* including the validity appraisal part (Chapters 3 to 6), and effect analyses (Chapters 7 and 8); and a *Discussion Section,* including summary and general discussion (Chapter 9).

This first introductory chapter provides a general background on NHs in the Netherlands, up-to-date depression figures and approaches to recognition and treatment of depressive features in NH residents. The AiD program for improving NH depression management is introduced, and, in closing, the aim and research questions of this thesis are outlined.

Area of research: residents with and without dementia in Dutch nursing homes

AiD is developed for NH residents. Due to demographical trends and age-related health problems, the need for long-term care services is increasing in industrialized countries.[1] In 2009, Dutch long-term care facilities had a capacity of more than 160.000 beds and provided care and services to more than 250.000 persons with different care dependency.[2] Residential and nursing home residents are mainly elderly with an average age of about 85 years.[3] Nursing home residents tend to be more care-dependent than those in residential homes where, originally, only basic nursing care has been provided. Dutch NH units provide 24-hour a day nursing care, assistance with daily living, and psychosocial, paramedical and personal care[1,4] AiD focuses primarily on NH residents.

AiD is a multidisciplinary program developed for residents of somatic and dementia special care units. Dutch nursing homes employ multidisciplinary teams which include nursing staff, an elderly care physician,[5] a psychologist, a recreational

* UKON is a collaboration between 12 care-organizations and the Department of Primary and Community Care of the Radboud University Nijmegen Medical Centre. See www.uko-n.nl for more information.

therapist, a physical therapist, and a speech therapist. Another characteristic of Dutch NHs is the presence of different types of nursing units. The AiD program is developed to be used by multidisciplinary care teams of two types of long term care units: so-called somatic units for residents with physical disabilities; and dementia special care (DSC) units for residents with psychogeriatric morbidity. Slightly more residents (55%) reside in DSC units than in somatic units.[3]

Depression is a serious problem in nursing homes

Depression is the main contributor to mental illnesses in nursing homes. Depression is a mood disorder characterized by a range of symptoms, such as depressed mood, loss of interest and sleep problems, which interfere significantly with ability to function normally. Depression is a major cause for persons being admitted to nursing home.[6,7] Recent data indicate that depression in newly admitted residents is more often present than dementia only, which makes depression the main contributor to the growing proportion of mental illnesses in NHs.[8] Although the prevalence of depression in residents may decline over time,[9] the condition persists in 45% to 63% of residents with depressive features[9,10] and 14-15% of residents without depression become depressed within twelve months.[10,11]

More depression in nursing homes than in general population. A recent meta-analysis of 11 studies on depression recognition in NHs showed an average prevalence of 28.8% (95% CI, 20.2-38.3%) for depression.[12] An earlier review of 39 studies on depression in NH residents showed an average prevalence of 15.5% for major depressive disorders, and 25.7% for minor depression.[13] These numbers can be considered higher than those in the general population. A review of 34 community-based studies showed 1.8% and 9.8% prevalence for later-life major and minor depression, respectively.[14]

Numerous risk factors exist for depression in NH populations. In general, somatic illness, functional impairment, cognitive impairment, lack of social support, and a history of depression are the main predictors for depression in elderly people.[15] These factors increase the likelihood of being admitted to a NH,[6,16,17] which implies that NH residents are at high risk of developing depression. In addition, loss of freedom and autonomy due to NH institutional regimen and regulations, lack of privacy, ever-present death and grief, and a lack of meaningful in-house activities are important factors contributing to depressive feelings in residents.[18]

Depression in NH residents amplifies other problems. Managing depression in NH residents is important because the condition may amplify other problems, such as

functional deficits, behavioral disturbances, non-compliance with treatment, and pain[19] Furthermore, depression is associated with increased mortality,[20-22] low quality of life,[23-25] increased risk of hospitalization and the use of other health care services.[26,27]

The high prevalence of depression, the numerous risk factors for depression in NH populations, and the fact that depression amplifies other problems call for early detection of depressive features and treatment in order to reduce emotional suffering.

Depression recognition in nursing homes needs improvement

Depression recognition is poor in nursing homes. Although recognition and treatment of depression constitutes a quality indicator in NHs[28] the condition often goes unnoticed and is undertreated.[29-33] Different reasons for poor depression recognition have been assumed, such as focus of the NH staff on other conditions in the daily care practice, and the fact that depressive symptoms can also be attributed to medical illness, functional impairments, and dementia.[28,34-36] NH staff needs to be well equipped to overcome difficulties in recognizing depression. Unfortunately, the contrary is the case and NH staff is poorly equipped to serve residents who need both psychiatric and nursing care.[37] Management of depression has not been protocolized in most Dutch NHs, and structural assessment using validated instruments has seldom been done.[38] To improve depression recognition, well-defined and easy-to-implement procedures are needed, describing which NH professionals should be involved, when, and how.

Involving nursing staff to improve depression recognition is important. First, it is of particular importance to involve nursing staff members that are engaged in the daily care of the resident. Practice nurses can be more sensitive in recognizing depression than NH physicians.[22,39] This can be due to the fact that nurses spend more time than physicians with patients.[12] However, those who have the most daily contact with the residents have the lowest level of knowledge of depression.[29] One of the important reasons for that is a lack of specific in-service training.[40] Nursing staff training may increase confidence in communicating depression-related information to physicians[41] and improve depression recognition.[42]

Language and memory impairments can impede reliable communication about depressive feelings. Second, to improve depression recognition, it is important to involve the unit psychologists and physician, who are trained in screening and diagnosing depression. The use of a validated depression instrument may help the psychologist to further evaluate the observations of the nursing staff, his or her own observations, and information from the resident. For this, it is important to consider that many NH residents,

mainly in DSC units, have language and memory impairments that impede reliable communication about their feelings. Therefore, the screening instruments should be validated in residents, both with and without these cognitive problems. Accordingly, the comprehensive diagnostic procedure that follows the screening procedure should also account for possible cognitive problems. Depressive symptoms in persons with dementia can be considered different from the symptoms of depression in persons without dementia.[43,44] However, the often used standardised criteria for depression[45,46] do not account for that. The use of diagnostic criteria that account for qualitative differences of depression in dementia and non-dementia is recommended.[47]

For improving depression recognition, AiD involves multiple disciplines and accounts for possible cognitive impairments of residents. The AiD care program equips the nursing staff, the psychologist and physician with tools to perform comprehensive depression assessment in residents of both DSC and somatic units. A training of 3.5 hours about depression constitutes the first step of the program implementation. For the purpose of involving the nursing staff, a new brief screening instrument was developed to be used on a regular basis in residents with and without dementia. Such regular use of a brief screening instrument constitutes the first phase of the AiD program (see Appendix 1). Furthermore, the AiD program provides the psychologists and physicians with evidence and practice-based guidelines on screening instruments that should be used in residents with and without cognitive problems (Phase 2), and guidelines on diagnostic criteria that should be used in persons with and without dementia (Phase 3).

Depression treatment: collaborative care is needed

Depression is merely treated with antidepressants despite scarce evidence of their effectiveness. A recently performed extensive search strategy in 9 databases showed that only two randomized controlled trials using a control group were performed on antidepressant medication in NH residents with depression; no benefit for medication over placebo was reported.[48] Furthermore, pharmacological treatment of depression in dementia may not be beneficial compared with placebo.[49,50] Yet, most NH residents with depression are treated merely with antidepressants.[38,51]

Evidence for psychosocial interventions is available. Bharucha et al.[52] reviewed psychosocial interventions in residents with and without dementia in long-term care facilities, including NHs. The authors found that a majority of 18 RCTs performed up to 2004 reported significant benefits on depression and various psychological well-being outcomes. Recent studies on non-pharmacological NH interventions imply that non-

pharmacological interventions including behavioral strategies can be effective.[53-56] Considering that adverse outcomes of drug treatment of depression can be serious in older people,[57] behavioral and psychosocial interventions should be considered as first-line treatment in NH residents.

Collaborative care of depression is more effective than a mono-therapy. A recent systematic review of reviews on treatment of depression showed that collaborative care and additional psychotherapy provides more benefits for patients than pharmacotherapy alone.[58] To date, only two studies in NHs are known on multidisciplinary collaborative care interventions with depression outcomes. Brodaty et al.[59] compared a formula-driven psychogeriatric team case management with a consultative approach, and usual care in NH residents with dementia. They found no difference between the three modes on depression outcomes. In another study, behavioral management in nursing and residential homes, training of caregivers by an old-age psychiatric hospital outreach team, and ongoing support to individual workers has shown positive effects on depression outcomes.[60] More research on multidisciplinary collaborative depression care in NHs is needed. Given the underrecognition of depression in NHs, adequate collaborative depression management comprising evidence-based collaborative treatment procedures should certainly include structural depression screening and diagnosing procedures (depression assessment).[61]

AiD provides protocols for collaborative depression treatment. The presence of multidisciplinary care teams in Dutch nursing homes offers ideal preconditions to manage depression effectively. Alongside the comprehensive assessment procedures aimed at the recognition and diagnosing depression, the AiD care program includes procedures for collaborative treatment (see Appendix 1). Protocols that enable the multidisciplinary teams to provide the best treatment to NH residents with depression or depressive symptoms were developed based on current Dutch and international guidelines, evidence from research on psychotherapy and pharmacotherapy, and existing clinical practice in Dutch NHs. The program provides protocols for nursing staff and the recreational therapist on behavioral activation strategies (Module 1); protocols for the psychologist on psychotherapy (Module 2); and a protocol for the physician on pharmacological therapy (Module 3). Although there is a certain degree of standardization in the program, treatment can be tailored according to the severity of depressive features and the cognitive status of the resident. The program accounts for the resident's ability to perform activities and to be involved in psychotherapy, and prescribes monitoring of treatment

results by the psychologists and the physician using depression-screening instruments from the assessment procedure.[†]

Apathy and depression

Monitoring apathy is important when treating depression. AiD behavioral strategies that activate residents with depressive features can also be beneficial for apathy,[62] whereas pharmacological treatment of depression using antidepressants may induce apathy.[63,64] Apathy refers to a low motivation disturbance evidenced by a lack of goal-directed behavior, cognition, and emotional affect.[65-67]

Given that apathy is associated with quality of life[68,69] it is important to monitor apathetic features when treating depression in NH residents. Monitoring apathy is especially indicated in dementia, because apathetic behavior is the most common non-cognitive disturbance in NH residents with dementia[70-72] and this disturbance is associated with a faster cognitive and functional decline[73] and increased mortality.[74,75] The AiD program does not include procedures for monitoring apathy. For the purpose of determining whether the program affected apathy, and for possible improvement of the program, apathy was monitored alongside the AiD trial.

Research questions

The general aim of this thesis is to evaluate the *validity* and the *effects* of the Act in case of Depression intervention. The first part of the Results section addresses the validity questions including validity of the instruments used in the program and in the trial. The effects of the program are presented in part two.

Validity appraisal

The first two research questions of this thesis address the validity of the instruments used for screening of depression in AiD. The third research question addresses the validity of an instrument used for monitoring apathy in the trial. The fourth question addresses the validity of the intervention.

Research question 1: Is the Nijmegen Observer-Rated Depression (NORD) scale a valid instrument for screening for depression in residents with and without dementia?

[†] Chapter 2 provides more detailed information about the program content. A schematic overview of the program pathways and abstracts of the program treatment protocols in English can be found in Apendix 1. The full texts in Dutch can be requested at the UKON website, www.uko-n.nl.

Implementation of new guidelines in the NH environment with limited staff time and numerous existing protocols can be difficult.[76] To facilitate recognition of depression in NH residents by involving nursing staff, the complexity of the procedures should be low.[77] A brief screening instrument that can be used in all residents regardless their language or cognitive status was preferred for AiD phase 1 recognition. A new observer-rated instrument was developed for this purpose because no brief instruments were developed for and validated in NH residents with and without cognitive problems. The results of a validity study on the accuracy of the NORD and its applicability in the NH population are presented in Chapter 3.

Research question 2: Is the Cornell Scale for Depression in Dementia (CSDD) an accurate proxy-based instrument in dementia when professional caregivers are the only available source of information?

The Geriatric Depression Scale used in phase 2 (see Appendix A1) was previously shown to be a valid instrument in NH residents without cognitive impairments[78,79] Yet, the CSDD was not validated in residents with severe cognitive problems when professional caregivers are the only source of information. Such validation is important considering that a large proportion of NH residents cannot be interviewed. Chapter 4 presents the results of a validation study in which the accuracy of the CSDD was tested and compared with the accuracy of another widely used depression scale. Given that nursing staff may experience problems in scoring of non-observable symptoms (e.g. suicidal thoughts), different methods to account for missing items were explored. The CSDD was also used in the trial to assess depression outcomes; therefore, the validation of the instrument is of importance for credibility of the effect results.

Research question 3: Is the abbreviated 10-item Apathy Evaluation Scale a valid instrument to assess apathy in nursing home residents?

The 10-item Apathy evaluation Scale (AES-10) was developed for NH population but has not yet been validated against diagnostic criteria for apathy. In order to determine whether the scale is useful in NH residents, a comparative validation of this scale, and another widely used instrument, was performed against recently proposed diagnostic criteria for apathy. The results are presented in Chapter 5. This instrument was also used for monitoring the effect of AiD on apathy and, therefore, its validity is of importance for credibility of the effect results.

Research question 4: Is AiD feasible in nursing homes, and can data of the AiD trial be used for analyzing the effectiveness of the program?

It is essential to evaluate the trial's sampling quality and intervention quality prior to executing statistical analyses on health intervention effects. Sampling quality addresses the recruitment and randomisation procedures. Evaluation of the sampling quality is important for both internal and external validity of the intervention. The intervention quality refers to, first, the feasibility of the program, which is important for the external validity appraisal. Second, intervention quality refers to the extent to which the program was performed, which is important for understanding both external and internal validity of the intervention. If the intervention does not possess internal validity, further statistical analyses could be meaningless due to lack of credibility of the results. If the intervention does not possess external validity, further implementation of the program in practice would be questioned. Process evaluation for the appraisal of sampling and intervention quality is presented in Chapter 6.

Intervention effects

The fifth and the sixth research questions address the effectiveness of the AiD care program.

Research question 5: Is the AID care program effective regarding depression prevalence?

In chapter 7, the results of the trial are presented regarding the main effects of AiD on depression prevalence. Secondary outcomes were the severity of depressive symptoms, and quality of life.

Research question 6: Does the AiD program influence apathy in NH residents?

Considering that the treatment of depression may influence apathy in NH residents, analyses were performed on the possible effect of AiD on apathy. For better insight into the effectiveness of the program, the effect of the AiD on different groups of depressive symptoms, which may or may not overlap with apathy, was analyzed. The results are presented in Chapter 8.

Literature

1. Ribbe MW, Ljunggren G, Steel K, et al. Nursing homes in 10 nations: a comparison between countries and settings. *Age and ageing* 1997;26 Suppl 2:3-12.
2. Statline C. AWBZ-zorg met verblijf; leeftijd en geslach: Verpleging en verzorging. http://statline.cbs.nl/StatWeb/publication. Retreaved on March 2, 2012. 2012.

3. Klerk Md. Zorg in de laatste jaren: Gezondheid en hulpgebruik in verzorgings- en verpleeghuizen 2000-2008.: Sociaal en Cultureel Planbureau, 2011.
4. Schols JM, Crebolder HF, van Weel C. Nursing home and nursing home physician: the Dutch experience. *J Am Med Dir Assoc* 2004;5:207-212.
5. Koopmans RTCM, Lavrijsen J, Hoek JF, et al. Dutch Elderly Care Physician: A new generation of nursing home physician specialists. *J Am Geriatr Soc* 2010;58:1807-1809.
6. Gaugler JE, Yu F, Krichbaum K, et al. Predictors of nursing home admission for persons with dementia. *Medical care* 2009;47:191-198.
7. Samus QM, Mayer L, Onyike CU, et al. Correlates of functional dependence among recently admitted assisted living residents with and without dementia. *J Am Med Dir Assoc* 2009;10:323-329.
8. Fullerton CA, McGuire TG, Feng Z, et al. Trends in mental health admissions to nursing homes, 1999-2005. *Psych Serv* 2009;60:965-971.
9. Smalbrugge M, Jongenelis L, Pot AM, et al. Incidence and outcome of depressive symptoms in nursing home patients in the Netherlands. *Am J Geriatr Psychiatry* 2006;14:1069-1076.
10. Barca ML, Engedal K, Laks J, et al. A 12 months follow-up study of depression among nursing-home patients in Norway. *J Affect Disord* 2010;120:141-148.
11. Boorsma M, Joling K, Dussel M, et al. The incidence of depression and its risk factors in Dutch nursing homes and residential care homes. *Am J Geriatr Psychiatry* In press.
12. Mitchell AJ, Kakkadasam V. Ability of nurses to identify depression in primary care, secondary care and nursing homes-A meta-analysis of routine clinical accuracy. *Int J Nurs Stud* 2011;48:359-358.
13. Jongenelis K, Pot AM, Eisses AM, et al. [Depression among older nursing home patients. A review]. *Tijdschr Gerontol Geriatr* 2003;34:52-59.
14. Beekman AT, Copeland JR, Prince MJ. Review of community prevalence of depression in later life. *Br J Psychiatry* 1999;174:307-311.
15. Djernes JK. Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatr Scand* 2006;113:372-387.
16. Thygesen E, Saevareid HI, Lindstrom TC, et al. Predicting needs for nursing home admission - does sense of coherence delay nursing home admission in care dependent older people? A longitudinal study. *Int J Older People Nurs* 2009;4:12-21.
17. Gaugler JE, Duval S, Anderson KA, et al. Predicting nursing home admission in the U.S: a meta-analysis. *BMC geriatrics* 2007;7:13.
18. Choi NG, Ransom S, Wyllie RJ. Depression in older nursing home residents: the influence of nursing home environmental stressors, coping, and acceptance of group and individual therapy. *Aging Ment Health* 2008;12:536-547.
19. Hyer L. Depression in Long Term Care. *Clin Psychol - Sci Pr* 2005;12:280-299.
20. Barca ML, Selbaek G, Laks J, et al. Factors associated with depression in Norwegian nursing homes. *Int J Geriatr Psychiatry* 2009;24:417-425.
21. Katz IR, Lesher E, Kleban M, et al. Clinical features of depression in the nursing home. *Int Psychogeriatr* 1989;1:5-15.
22. Rovner BW, German PS, Brant LJ, et al. Depression and mortality in nursing homes. *JAMA* 1991;265:993-996.
23. Hoe J, Hancock G, Livingston G, et al. Quality of life of people with dementia in residential care homes. *Br J Psychiatry* 2006;188:460-464.
24. Wetzels RB, Zuidema SU, de Jonghe JF, et al. Determinants of quality of life in nursing home residents with dementia. *Dement Geriatr Cogn Disord* 2010;29:189-197.

25. Gonzalez-Salvador T, Lyketsos CG, Baker A, et al. Quality of life in dementia patients in long-term care. *Int J Geriatr Psychiatry* 2000;15:181-189.
26. Smalbrugge M, Pot AM, Jongenelis L, et al. The impact of depression and anxiety on well being, disability and use of health care services in nursing home patients. *Int J Geriatr Psychiatry* 2006;21:325-332.
27. Bartels SJ, Horn SD, Smout RJ, et al. Agitation and depression in frail nursing home elderly patients with dementia: treatment characteristics and service use. *Am J Geriatr Psychiatry* 2003;11:231-238.
28. AGS/AAGP. The American Geriatrics Society and American Association for Geriatric Psychiatry recommendations for policies in support of quality mental health care in U.S. nursing homes. J Am Geriatr Soc 2003;51:1299-1304.
29. Ayalon L, Arean P, Bornfeld H. Correlates of knowledge and beliefs about depression among long-term care staff. *Int J Geriatr Psychiatry* 2008;23:356-363.
30. Gruber-Baldini AL, Zimmerman S, Boustani M, et al. Characteristics associated with depression in long-term care residents with dementia. *Gerontologist* 2005;45:50-55.
31. Kramer D, Allgaier AK, Fejtkova S, et al. Depression in nursing homes: prevalence, recognition, and treatment. *Int J Psychiatry Med* 2009;39:345-358.
32. Snowdon J, Fleming R. Recognising depression in residential facilities: An Australian challenge. *Int J Geriatr Psychiatry* 2008;23:295-300.
33. Damian J, Pastor-Barriuso R, Valderrama-Gama E. Descriptive epidemiology of undetected depression in institutionalized older people. *J Am Med Dir Assoc* 2010;11:312-319.
34. Thakur M, Blazer DG. Depression in long-term care. *J Am Med Dir Assoc* 2008;9:82-87.
35. Kurlowicz LH, Evans LK, Strumpf NE, et al. A psychometric evaluation of the Cornell Scale for Depression in Dementia in a frail, nursing home population. *Am J Geriatr Psychiatry* 2002;10:600-608.
36. McCabe MP, Davison T, Mellor D, et al. Barriers to care for depressed older people: perceptions of aged care among medical professionals. *Int J Aging Hum Dev* 2009;68:53-64.
37. Grabowski DC, Aschbrenner KA, Rome VF, et al. Quality of mental health care for nursing home residents: a literature review. *Med Care Res Rev* 2010;67:627-656.
38. Dorland L, Pot A, Verbeek M, et al. In Dutch: [Psychische hulpverlening voor ouderen in verzorgings- en verpleeghuizen; deelstudie 7.] In Monitor geestelijke gezondheidszorg ouderen, rapportage 2006 Utrecht: Trimbos-Instituut, 2007.
39. Falck RP, Pot AM, Braam AW, et al. [Prevalence and diagnosis of depression in frail nursing home patients; a pilot study]. *Tijd Geront & Geriatr* 1999;30:193-199.
40. Bagley H, Cordingley L, Burns A, et al. Recognition of depression by staff in nursing and residential homes. *Journal of clinical nursing* 2000;9:445-450.
41. Brown EL, Raue PJ, Roos BA, et al. Training nursing staff to recognize depression in home healthcare. *J Am Geriatr Soc* 2010;58:122-128.
42. Eisses AM, Kluiter H, Jongenelis K, et al. Care staff training in detection of depression in residential homes for the elderly: randomised trial. *Br J Psychiatry* 2005;186:404-409.
43. Barca ML, Selbaek G, Laks J, et al. The pattern of depressive symptoms and factor analysis of the Cornell Scale among patients in Norwegian nursing homes. *Int J Geriatr Psychiatry* 2008;23:1058-1065.
44. Olin JT, Katz IR, Meyers BS, et al. Provisional diagnostic criteria for depression of Alzheimer disease: rationale and background. *Am J Geriatr Psychiatry* 2002;10:129-141.
45. APA. Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association, 2000.

46. International Statistical Classification of Diseases, 10th Revision (ICD-10) Geneva, Switzerland: World Health Organization, 1992.
47. Koopmans R, Zuidema S, Leontjevas R, et al. Comprehensive assessment of depression and behavioral problems in long-term care. *Int Psychogeriatr* 2010;22:1054-1062.
48. Boyce RD, Hanlon JT, Karp JF, et al. A Review of the Effectiveness of Antidepressant Medications for Depressed Nursing Home Residents. *J Am Med Dir Assoc* 2011.
49. Nelson JC, Devanand DP. A systematic review and meta-analysis of placebo-controlled antidepressant studies in people with depression and dementia. *J Am Geriatr Soc* 2011;59:577-585.
50. Banerjee S, Hellier J, Dewey M, et al. Sertraline or mirtazapine for depression in dementia (HTA-SADD): a randomised, multicentre, double-blind, placebo-controlled trial. *Lancet* 2011;378:403-411.
51. Levin CA, Wei W, Akincigil A, et al. Prevalence and treatment of diagnosed depression among elderly nursing home residents in Ohio. *J Am Med Dir Assoc* 2007;8:585-594.
52. Bharucha AJ, Dew MA, Miller MD, et al. Psychotherapy in long-term care: A review. *J Am Med Dir Assoc* 2006;7:568-580.
53. Konnert C, Dobson K, Stelmach L. The prevention of depression in nursing home residents: a randomized clinical trial of cognitive-behavioral therapy. *Aging Ment Health* 2009;13:288-299.
54. Hyer L, Yeager CA, Hilton N, et al. Group, individual, and staff therapy: an efficient and effective cognitive behavioral therapy in long-term care. *Am J Alzheimers Dis Other Demen* 2008;23:528-539.
55. Buettner LL, Fitzsimmons S. AD-venture program: therapeutic biking for the treatment of depression in long-term care residents with dementia. *Am J Alzheimers Dis Other Demen* 2002;17:121-127.
56. Verkaik R, Francke AL, van Meijel B, et al. The effects of a nursing guideline on depression in psychogeriatric nursing home residents with dementia. *Int J Geriatr Psychiatry* 2011;26:723-732.
57. Coupland C, Dhiman P, Morriss R, et al. Antidepressant use and risk of adverse outcomes in older people: population based cohort study. *BMJ* 2011;343:d4551.
58. Oestergaard S, Moldrup C. Improving outcomes for patients with depression by enhancing antidepressant therapy with non-pharmacological interventions: a systematic review of reviews. *Public health* 2011;125:357-367.
59. Brodaty H, Draper BM, Millar J, et al. Randomized controlled trial of different models of care for nursing home residents with dementia complicated by depression or psychosis. *J Clin Psychiat* 2003;64:63-72.
60. Proctor R, Burns A, Powell HS, et al. Behavioural management in nursing and residential homes: a randomised controlled trial. *Lancet* 1999;354:26-29.
61. Psychiatry AGSaAAfG. Consensus statement on improving the quality of mental health care in U.S. nursing homes: management of depression and behavioral symptoms associated with dementia. *J Am Geriatr Soc* 2003;51:1287-1298.
62. Brodaty H, Burns K. Nonpharmacological Management of Apathy in Dementia: A Systematic Review. *Am J Geriatr Psychiatry* 2012;20(7):549-564.
63. Barnhart WJ, Makela EH, Latocha MJ. SSRI-induced apathy syndrome: a clinical review. *J Psychiatr Pract* 2004;10:196-199.
64. Settle EC, Jr. Antidepressant drugs: disturbing and potentially dangerous adverse effects. *The Journal of clinical psychiatry* 1998;59 Suppl 16:25-30; discussion 40-22.
65. Marin RS. Apathy: a neuropsychiatric syndrome. *J Neuropsychiatry Clin Neurosci* 1991;3:243-254.
66. Starkstein SE, Petracca G, Chemerinski E, et al. Syndromic validity of apathy in Alzheimer's disease. *Am J Psychiatry* 2001;158:872-877.

67. Robert P, Onyike CU, Leentjens AF, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatr* 2009;24:98-104.
68. Samus QM, Rosenblatt A, Steele C, et al. The association of neuropsychiatric symptoms and environment with quality of life in assisted living residents with dementia. *Gerontologist* 2005;45 Spec No 1:19-26.
69. Gerritsen DL, Jongenelis K, Steverink N, et al. Down and drowsy? Do apathetic nursing home residents experience low quality of life? *Aging Ment Health* 2005;9:135-141.
70. Aalten P, de Vugt ME, Jaspers N, et al. The course of neuropsychiatric symptoms in dementia. Part I: findings from the two-year longitudinal Maasbed study. *Int J Geriatr Psychiatry* 2005;20:523-530.
71. Zuidema SU, Derksen E, Verhey FR, et al. Prevalence of neuropsychiatric symptoms in a large sample of Dutch nursing home patients with dementia. *Int J Geriatr Psychiatry* 2007;22:632-638.
72. Lyketsos CG, Olin J. Depression in Alzheimer's disease: overview and treatment. *Biol Psychiatry* 2002;52:243-252.
73. Starkstein SE, Jorge R, Mizrahi R, et al. A prospective longitudinal study of apathy in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2006;77:8-11.
74. Diesfeldt HF, van Houte LR, Moerkens RM. Duration of survival in senile dementia. *Acta Psychiat Scand* 1986;73:366-371.
75. Van Dijk PT, Dippel DW, Habbema JD. A behavioral rating scale as a predictor for survival of demented nursing home patients. *Arch Gerontol Geriatr* 1994;18:101-113.
76. Colon-Emeric CS, Lekan D, Utley-Smith Q, et al. Barriers to and facilitators of clinical practice guideline use in nursing homes. *J Am Geriatr Soc* 2007;55:1404-1409.
77. Grol RP, Wensing M. *Implementatie: Effectieve verbetering van de patiëntenzorg*. Amsterdam: Elsevier gezondheidszorg, 2010.
78. Jongenelis K, Gerritsen DL, Pot AM, et al. Construction and validation of a patient- and user-friendly nursing home version of the Geriatric Depression Scale. *Int J Geriatr Psychiatry* 2007;22:837-842.
79. Yesavage JA, Brinka TL, Roseb TL, et al. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* 1983;17:37-49.
80. Leontjevas R, Gerritsen DL, Vernooij-Dassen MJFJ, et al. The Nijmegen Observer-Rated Depression Scale for Detection of Depression in Nursing Home Residents. *Int J Geriatr Psychiatr* In press.
81. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23:271-284.
82. Olin JT, Schneider LS, Katz IR, et al. Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002;10:125-128.

Chapter 2

Act in case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes

Study protocol

Debby L. Gerritsen, Martin Smalbrugge, Steven Teerenstra, Ruslan Leontjevas, Eddy M. Adang, Myrra J.F.J. Vernooij-Dassen, Els Derksen, Raymond T.C.M. Koopmans

BMC Psychiatry 2011, 11:91

http://www.biomedcentral.com/1471-244X/11/91

Abstract

Background: The aim of this study is evaluating the (cost-) effectiveness of a multidisciplinary, evidence based care program to improve the management of depression in nursing home residents of somatic and dementia special care units. The care program is an evidence based standardization of the management of depression, including standardized use of measurement instruments and diagnostical methods, and protocolized psychosocial, psychological and pharmacological treatment.

Methods/Design: In a 19-month longitudinal controlled study using a stepped wedge design, 14 somatic and 14 dementia special care units will implement the care program. All residents who give informed consent on the participating units will be included. Primary outcomes are the frequency of depression on the units and quality of life of residents on the units. The effect of the care program will be estimated using multilevel regression analysis. Secondary outcomes include accuracy of depression-detection in usual care, prevalence of depression-diagnosis in the intervention group, and response to treatment of depressed residents. An economic evaluation from a health care perspective will also be carried out.

Discussion: The care program is expected to be effective in reducing the frequency of depression and in increasing the quality of life of residents. The study will further provide insight in the cost-effectiveness of the care program.

Trial registration: Netherlands Trial Register (NTR): NTR1477

Background

Depression is a common health problem in nursing home (NH) residents: prevalence rates vary from 6 to even 50%.[1-3] Depression is strongly related to quality of life of NH residents,[4] it seriously impacts wellbeing and daily functioning, and increases use of health care services and even mortality.[5-8] The association between depression and quality of life highlights the importance of identifying and treating depression in NH residents with and those without dementia.[4,9] Unfortunately, although depression is a treatable disorder, 9 various studies have shown poor detection and undertreatment of depression in NH residents.[2,10-12]

Several studies have demonstrated effects of pharmacological and psychosocial interventions for depression in nursing homes.[13,14] The review of Bharucha et al.[15] of 'talk therapies' for depression in long-term care presents evidence for an improvement in depressive symptoms after reminiscence/life review therapy. Moreover, there is evidence for the effectiveness of multifaceted interventions in residential care[16-18] and in nursing homes.[19,20]

The Nijmegen University NH Network (UKON), a collaboration between 12 care organizations and the Department of Primary and Community Care of the Radboud University Nijmegen Medical Centre, has developed the care program Act in case of Depression (AiD), a multidisciplinary care program to identify and treat depression and monitor treatment effects. The care program is based on and in accordance with the recommendations as formulated in the Supplement Older Adults of the multidisciplinary evidence based guideline for diagnosis and treatment of depression[21] and the Consensus Statement of the American Geriatrics Society and the American Association of Geriatric Psychiatry.[22] The care program is an implementable plan of work that coordinates how the different disciplines should work together, fits in daily practice, and describes how new working methods are related to and can be integrated in the present care process following a step-by-step plan.[23]

To date, cost effectiveness studies into the management of depression in NH have not been carried out, but are requested.[24] Gruber-Baldini et al.[10] did find increased involvement of mental health professionals in depressed long-term care residents with dementia, and Smalbrugge et al.[6] found that depressed residents of somatic units had increased use of medication, and received medical specialist consultation and treatment more often than non-depressed residents, implying expensive medical tests and hospital

admissions. This paper describes a study that will evaluate the cost effectiveness of the care program AID.

Methods/Design

The study is a stepped wedge, multicentre intervention study on 14 somatic and 14 Dementia Special Care (DSC) units of UKON.

A stepped wedge design is a type of crossover design in which different clusters (here: units) cross over from the control group to the intervention group at different time points. All clusters are measured at each time point. The first time point corresponds to a baseline measurement where none of the clusters receive the intervention of interest; at the last time point all clusters receive the intervention. After intermediate time points, clusters initiate the intervention. More than one cluster may start the intervention at a time point, but the time a cluster begins the intervention is randomized[25] (see Figure 2.1 for a graphical representation of the design). This way, comparisons within units ánd between units will be available, making the design very powerful. Another advantage of the design is that all involved units will receive the intervention – which is expected to increase motivation for participating in the study.

		Measurement					
		T0	T1	T2	T3	T4	T5
Group	1	0	1	1	1	1	1
	2	0	0	1	1	1	1
	3	0	0	0	1	1	1
	4	0	0	0	0	1	1
	5	0	0	0	0	0	1

Figure 2.1. Graphical representation of the stepped wedge design

'0' represents measurement of usual care; control condition; '1' represents measurement after the intervention has been implemented; intervention-condition

At the start of the data collection, the residents with informed consent of all 28 units are screened for depression (T0). Following this, each of the units is randomly assigned to one of 5 groups. Each group starts the intervention at different time points, directly after one

of the measurements (T0-T4), which are each 4 months apart. In the four- month interval between T0 and T1, nursing staff of the first group is trained within the first month. After this month, the intervention runs for the subsequent 3 months in the first group before the second measurement (T1) of all 28 units takes place. After T1, the second group is trained, and the intervention starts in this group while it is continued in the first group. This procedure is repeated for the remaining 3 groups until, at the last measurement (T5), all 28 units are in the intervention condition. Consequently, the follow up in the intervention condition varies from 3 months for the last group, which starts with the intervention 1 month after T4, to 19 months for the first group, which starts after T0.

Intervention

Figure 2.2 offers a graphical representation of the care program AID. AID proposes an evidence and practice based standardization of 5 components in the management of depression: 1) identification of depressive symptoms, 2) screening, 3) diagnosis, 4) treatment and 5) monitoring. AID includes standardized use of measurement instruments and diagnostical methods, and protocolized treatment that combines psychosocial, psychological and pharmacological interventions. Cooperation between the disciplines is prearranged. As the ability of nursing staff to detect depression can and should be enhanced,[26] the multifaceted and multidisciplinary care program 'AID' starts with a training program for nursing staff on how to identify symptoms of depression using a short observation scale[27] and how to support NH residents with depressive symptoms or depression. Further, AID comprises plans of work for the identification, screening, diagnosing, treatment and monitoring of depression.

Identification (or detection): Nursing staff completes a short observation scale for depression[27] for all participating residents on the unit. If according to the scores on the scale further screening is indicated, nursing staff contacts the psychologist who takes over the coordination on the screening and diagnosing. If no further screening is indicated, nursing staff will complete the observation scale again after 3-4 months.

Screening: The psychologist screens the 'identified' residents of somatic units for depressive symptoms with the GDS8 (Geriatric Depression Scale-NH version; cut-off score 2/3)[28] and those of DSC units with the CSDD (Cornell Scale for Depression in Dementia; cut-off score 7/8).[29-31] If screening with the GDS8 in somatic residents is problematic because of cognitive or communication problems, the CSDD will also be administered.[32]

For residents with depressive symptoms, i.e. total scores on the GDS8 or CSDD above the cut-off score, a diagnostic procedure will follow. For other residents, the identification phase will be repeated after 3-4 months.

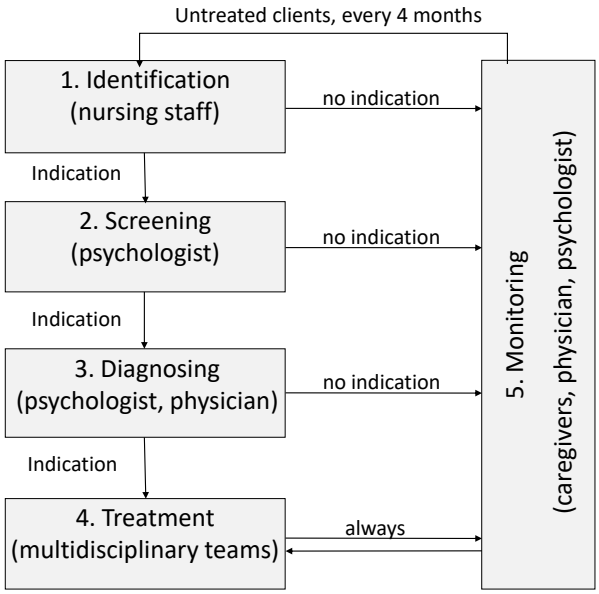


Figure 2.2 Graphical representation of the AID care program

Diagnosing: The elderly care physician and psychologist of each unit perform a diagnostic procedure including the use of chart information, caregiver interview, and examination of the resident (interview, physical examination). Diagnosis of major depression is established according to the DSM-IV-TR criteria. For minor depression the same criteria are used while only 2 to 4 symptoms are present.[33,34] In residents with dementia the Provisional Diagnostic Criteria for depression in Alzheimer's disease are applied (PDC).[35]

Treatment: **Somatic and dementia residents with depressive symptoms,** but without a clinical diagnosis of depression, are offered a personal day structure program made by the nursing staff in collaboration with the recreational therapist. Exercise and music therapy can be part of this day program. Psycho-education is also offered to the resident and/or relatives, including information about depressive symptoms and coping strategies.

Somatic residents with minor depression receive the same treatment as residents with depressive symptoms extended with individual life review therapy. This therapy is based on a protocol that has already been used successfully in Dutch residential care residents and is developed in close collaboration with the Dutch life review expert E. Bohlmeijer.[36]

Somatic residents with major depressive disorder receive the same treatment as residents with minor depression extended with pharmacological treatment, when deemed appropriate by the elderly care physician. Prescription of pharmacological therapy is in accordance with the recommendations of the Supplement Older Adults.[21]

For dementia residents with a PDC-depression diagnosis, treatment includes a personal day structure program, a behavioral management strategy developed by the psychologist and psycho education - especially of relatives. Apart from that, psychological treatment is offered: the clinical experts involved in the development of this care program agreed with recommendations made in the Supplement Older Adults[21] to intervene through the nursing staff (mediative therapy), but stressed that individual contact with the resident is also a necessity. Thus, for dementia residents, psychological treatment comprises of the psychologist supporting and supervising the nursing staff and recreational therapist more intensively in their execution of the day structure program and behavioral management strategy. This support takes place in a regular staff meeting, every two weeks. Within 1 month after the diagnosis, the day structure program and behavioral management strategy should be incorporated in regular care. The psychologist supervises the recreational therapist and nursing staff in at least 2 regular staff meetings. Additionally, if the depression in dementia residents is severe, pharmacological therapy can be given by the elderly care physician, when deemed appropriate.

Monitoring with a validated measurement instrument takes place to evaluate treatment. For this purpose, the GDS8 is used in somatic residents, and the CSDD is used in dementia residents.

Sampling

We calculated the sample-size using the following assumptions. For somatic units: 25 residents per unit,[37] a depression prevalence of 22%,[38] a remission rate of 40%,[39] and an attrition of 20%.[38] For DSC units: 20 residents per unit,[40] a depression prevalence of 30%,[10,40-42] a remission rate of 35%,[13] and negligible attrition.[40] Based on these assumptions and a significance level alpha of 0.05, a power of 0.80 and an ICC of 0.1 for both somatic and dementia residents, 16 clusters (units) with 6 measurements are needed in a stepped wedge design to allow multilevel analysis.

Given that the outcomes will be presented on unit level, during the data collection, newly admitted residents and/or their legal representatives are asked to provide informed consent on all units. This way, the sample size is not influenced by death or relocation of participating residents and can remain stable.

Ethical approval

The Medical Ethics Committee of the Radboud University Nijmegen Medical Centre (CMO Arnhem-Nijmegen) rated the study and pronounced that it is not burdensome for the participant. Each NH resident and/or the legal representative on the participating units receives written and verbal information prior to the AID study and is only included in the study after having given written informed consent.

Measurements

Primary outcomes are frequency of depression and quality of life. Frequency of depression (the percentage of residents with depression on a unit) is measured in somatic residents by a shortened version of the Geriatric Depression Scale (GDS),[43] the 8-item GDS-nursing home version (GDS8) of Jongenelis et al.,[28] which was made by deleting GDS-items that are not applicable to most NH residents. The GDS8 was validated in the AGED dataset, where it showed a good internal consistency of α=0.80 and high sensitivity rates of 96.3% for major depression and 83.0% for minor depression, with a specificity rate of 71.7% at a cut-off score of 3 or more.[28] The GDS8 also appears to be able to assess (change in) severity of depression.[44] The GDS-30 is originally a self-report instrument, the GDS8 is interview based. Frequency of depression in dementia residents is measured by the Cornell Scale for Depression in Dementia (CSDD).[29] The CSDD is administered through interviewing nursing staff about their observations of the residents' behavior. The CSDD consists of 19 items each rated as 0=absent, 1=mild or intermittent and 2=severe. The scores of the individual items are summed and a cut-off of 8 or more indicates depression.[29] Vida et al.[30] reported for a cut-off score of 8 or more, a sensitivity of 90% and specificity of 75% in residents with Alzheimer's Disease.

Quality of life in somatic residents is measured by the EQ-5D.[45] The EQ-5D instrument is a standardized non disease-specific instrument for describing and valuing Health Related Quality of Life.[46] There are two core components of the instrument: a description of the respondent's 'own health' using a health state classification system with five domains (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a rating of 'own general health' by means of a visual analogue 'thermometer' scale. The EQ5D has shown a good validity and good test-retest reliability.[47,48] In dementia residents quality of

life is measured by the EQ-5D proxy version.[49] Thereto, nursing staff are asked to score the scale for the resident.

Secondary outcomes are percentage accuracy of depression-detection in usual care, prevalence of depression-diagnosis in the intervention group, and response to treatment of depressed residents.

Additional measurements involve measurement of cognitive functioning by the Mini Mental State Examination (MMSE)[50] and measurement of sociodemographic variables, mental health history (including prior depressive episodes), present mental health condition (including a dementia diagnosis), possible treatment for depression, and somatic comorbidity.

Measurements are done by the research team. To study the compliance to the care program, the actual use of all components of the psychosocial, psychological and pharmacological treatment, as well as the factors determining this use, are registered. Accordingly, written checklists are used for nursing staff, recreational therapist, psychologist and elderly care physician, separately.

Data-analysis

Primary effects will be calculated using multilevel regression analysis, for somatic and DSC units separately. The GDS8-scores and CSDD-scores will be used in the primary analysis. Age, sex, cultural background and cognitive status will be used as covariates. The EQ5D will be analyzed as another primary outcome in the intervention study. For cost analysis, see economic evaluation. A process analysis will be carried out to determine the actual use of the components of the psychosocial, psychological and pharmacological treatment, and to determine facilitators and obstacles.

Secondary outcomes (percentage accuracy of depression-detection in usual care, prevalence of depression-diagnosis in the intervention group and response to treatment of depressed residents) will be analyzed using descriptive statistics.

Economic evaluation

This study investigates the efficiency of the care program AID compared to usual care as provided in NH units. If the program AID turns out to be successful, a decrease in the prevalence of depression in NH will occur. On the one hand the program needs investment in for example training of nursing staff and, consequently, generates extra costs compared to usual care. On the other hand it potentially generates savings as it reduces depression related time investment in NH.

The economic evaluation is based on the general principles of a cost-effectiveness analysis from a healthcare viewpoint. Based on the above mentioned primary outcomes, two different incremental cost effectiveness ratios (ICERs) will be computed, answering the questions: 'How much money has to be invested additionally in the care program to gain one percentage point decrease in frequency of depression?' and 'How much money has to be invested additionally in the care program to gain one Quality-Adjusted Life Year (QALY)?'

The cost analysis consists of two main parts. First, on resident level, volumes of care (to determine the incremental direct health care costs) based on the production process of the care program and of depression decrease are measured prospectively using an activity based costing approach. Focusing on activities performed with costs accumulated at the activity level(s) of the health care production processes, standardized case report forms are used to assess time invested by nursing staff, psychologist, elderly care physician and recreational therapist. Also, number of hospital admissions (number of days in hospital) and use of antidepressant medication are recorded.

Second, the cost prices for each volume of consumption will be determined to use these for multiplying the volumes registered for each participating resident. The Dutch guidelines for cost analyses will be used.[51] For units of care/resources where no guideline or standard prices are available, real cost prices will be determined. Statistics of the total costs per resident will be determined for usual care and care according to the care program AID. Depending on the skewness of the parameter distributions, statistical testing of differences between strategies will be of a parametrical or non-parametrical nature. The impact of deterministic variables, such as cost prices for volume parameters that are incremental cost drivers will be investigated using sensitivity analyses on the basis of the range of extremes.

The effect analysis adheres to the design of the study. Relevant for the economic evaluation are the frequency of depression (measured with GDS8 and CSDD) and QALYs (utilities measured with the EQ-5D). Using the trapezium rule, the QALYs will be computed in order to perform a cost-effectiveness analysis comparing the two alternative strategies. Change in utilities (EQ-5D) will be based on the mean values for the residents when they are in the control condition and the mean values after having been in the intervention for 3 (all 5 groups), 7 (4 groups), 11 (3 groups), 15 (2 groups) and 19 months (1 group). ICERs will be computed and sampling uncertainty will be determined using the bootstrap or Fieller method. Finally, a cost-effectiveness acceptability curve will be derived that is able to evaluate efficiency by different thresholds for the ICERs.

Discussion

In this paper we described the design of a randomized trial to evaluate the (cost) effectiveness of a multidisciplinary, evidence based care program to improve the management of depression in NH residents of somatic and DSC units. This study holds several unique elements.

First of all, the Department of Primary and Community Medicine of the Radboud University Nijmegen Medical Centre has established a structural collaboration with 12 care organizations (representing 40 NH and 100 residential homes) in the Nijmegen University NH Network (UKON). An expert group of the UKON has developed the care program AID, based on evidence based guidelines and the Consensus Statements.[21,22] Implementation is expected to be successful, because it fits with daily practice and describes how new working methods are related to and can be integrated in the present care process following a step-by-step plan.[23]

Secondly, the intervention is based on a stepped care approach: the more serious the depressive complaints or the depression, the more intense the intervention will be. The standardized interventions will be tailored to the needs of the individual resident. This will expectedly increase its effectiveness and facilitate transferring this strategy to other nursing homes.

Finally, the design of the study -the stepped wedge design- is a relatively new design, and has not been applied before in long term care. Using a stepped wedge design signifies that all participating units will cross-over from the control condition to the intervention condition during the study. This is expected to increase the motivation of NH workers to participate in scientific research.

In conclusion, the care program is expected to be effective in reducing the frequency of depression and in increasing the quality of life of residents. The study also will provide insight in the program's cost- effectiveness.

Literature

1. Jongenelis K, Pot AM, Eisses AM, Beekman AT, Kluiter H, van Tilburg W, Ribbe MW: Depression among older nursing home patients: A review. *Tijdschr Gerontol Geriatr* 2003, 34:52-59.
2. Kramer D, Allgaier AK, Fejtkova S, Mergl R, Hegerl U: Depression in nursing homes: prevalence, recognition, and treatment. *Int J Psychiatry Med* 2009, 39:345-358.
3. Snowdon J, Rosengren D, Daniel F, Suyasa L: Australia's use of the Cornell scale to screen for depression in nursing homes. *Australas J Ageing* 2011, 30:33-36.
4. Hoe J, Hancock G, Livingston G, Orrell M: Quality of life of people with dementia in residential care homes. *Br J Psychiatry* 2006, 188:460-464.
5. Demyttenaere K, Hemels MEH, Hudry J, Annemans L: A cost-effictiveness model of Escitalopram, Citalopram and Venlafaxine as first-line treatment for major depressive disorder in Belgium. *Clin Therapeutics* 2005, 27:111-124.
6. Smalbrugge M, Pot AM, Jongenelis K, Gundy CM, Beekman AT, Eefsting JA: The impact of depression and anxiety on well being, disability and use of health care services in nursing home patients. *Int J Geriatr Psychiatry* 2006, 21:325-332.
7. Beekman ATF, Deeg DJH, Braam AW, Smit JH, van Tilburg W: Consequences of major and minor depression in later life: a study of disability, well-being and service utilization. *Psychol Med* 1997, 27:1397-1409.
8. Rovner BW: Depression and increased risk of mortality in the nursing home patient. *Am J Med* 1993, 94(suppl 5A): 19S-23S.
9. Selwood A, Thorgrimsen L, Orrell M: Quality of life in dementia--a one-year follow-up study. *Int J Geriatr Psychiatry* 2005, 20:232-237.
10. Gruber-Baldini AL, Zimmerman S, Boustani M, Watson LC, Williams CS, Reed PS: Characteristics associated with depression in long-term care residents with dementia. Gerontologist 2005, 45SI1:50-55.
11. Jongenelis K: Depression in Dutch nursing homes: The Amsterdam Groningen Elderly Depression Study. PhD Thesis. VU University Medical Center; 2006.
12. Ayalon L, Arean P, Bornfeld H: Correlates of knowledge and beliefs about depression among long-term care staff. *Int J Geriatr Psychiatry* 2008, 23:356-363.
13. Snowden M, Sato K, Roy-Byrne P: Assessment and treatment of nursing home residents with depression or behavioral symptoms associated with dementia: a review of the literature. *J Am Geriatr Soc* 2003, 51:1305-1317.
14. Peskind ER: Management of depression in long-term care of patients with Alzheimer's disease. *J Am Med Dir Assoc* 2003, 4:S141-145.
15. Bharucha AJ, Dew MA, Miller MD, Borson S, Reynolds C: Psychotherapy in long-term care: a review. *J Am Med Dir Assoc* 2006, 7:568-580.
16. Llewellyn-Jones RH, Baikie KA, Smithers H, Cohen J, Snowdon J, Tennant CC: Multifaceted shared care intervention for late life depression in residential care: randomised controlled trial. *BMJ* 1999, 319:676-682.
17. Lewellyn-Jones RH, Baikie KA, Smithers H, Cohen J, Snowdon J, Tennant CC : How to help depressed older people living in residential care: a multifaceted shared-care intervention for late-life depression. *Int Psychogeriatr* 2001, 13:477-492.
18. Cuijpers P, van Lammeren P: Secondary prevention of depressive symptoms in elderly inhabitants of residential homes. *Int J Geriatr Psychiatry* 2001, 16:702-708.
19. Llewellyn-Jones RH, Snowdon J: Depression in nursing homes: ensuring adequate treatment. *CNS Drugs* 2007, 21:627-640.

20. Brooker DJ, Woolley RJ, Lee D: Enriching opportunities for people living with dementia in nursing homes: An evaluation of a multi-level activity-based model of care. *Aging Ment Health* 2007, 11:361-370.
21. Kok R: Multidisciplinaire richtlijn depressie, addendum ouderen. Utrecht: CBO/ Trimbosinstituut; 2008.
22. American Geriatrics Society & American Association for Geriatric Psychiatry: Consensus statement on improving the quality of mental health care in US nursing homes: management of depression and behavioural symptoms associated with dementia. *J Am Geriatr Soc* 2003, 51:1287-1298.
23. Grol RP, Wensing M (Eds): Implementatie: Effectieve verbetering van de patiëntenzorg. Amsterdam: Elsevier gezondheidszorg; 2010.
24. Lyne KJ, Barrett P, Evans C, Barkham M: Dimensions of variation on the CORE-OM. *Br J Clin Psychol* 2006, 45(Pt 2):185-203.
25. Hussey M, Hughes JP: Design and analysis of stepped wedge cluster randomized trials. *Contemp Clinical Trials* 2007, 28:182-191.
26. Eisses AM, Kluiter H, Jongenelis K, Pot AM, Beekman AT, Ormel J: Care staff training in detection of depression in residential homes for the elderly: randomised trial. *Br J Psychiatry* 2005, 186:404-409.
27. Hammond MF, O'Keeffe ST, Barer DH: Development and validation of a brief observer-rated screening scale for depression in elderly medical patients. *Age Ageing* 2000, 29: 511-515.
28. Jongenelis K, Gerritsen DL, Pot AM, Beekman AT, Eisses AM, Kluiter H, Ribbe MW: Construction and validation of a patient and user-friendly nursing home version of the Geriatric Depression Scale. *Int J Geriatr Psychiatry* 2007, 22:837-842.
29. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA: Cornell Scale for Depression in Dementia. *Biol Psychiatry* 1988; 23:271-284.
30. Vida S, Des Rosiers P, Carrier L, Gauthier S: Depression in Alzheimer's disease: receiver operating characteristic analysis of the Cornell Scale for Depression in Dementia and the Hamilton Depression Scale. *J Geriatr Psychiatry Neurol* 1994, 7:159-162.
31. Kørner A, Lauritzen L, Abelskov K, Gulmann N, Brodersen AM, Wedervang-Jensen T, Kjeldgaard KM: The Geriatric Depression Scale and the Cornell Scale for Depression in Dementia: A validity study. *Nord J Psychiatry* 2006, 60: 360-364.
32. UVNN, Nieuwboer M, Derksen E: Meetschalen stemming in het verpleeghuis: verslag van consensustraject stemming. Nijmegen; 2007.
33. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington DC; 2000.
34. Koopmans RT, Lavrijsen JC, Hoek JF, Went PB, Schols JM: Dutch elderly care physician: a new generation of nursing home physician specialists. *J Am Geriatr Soc* 2010, 58:1807-1809.
35. Olin JT, Katz IR, Meyers BS, Schneider LS, Lebowitz BD: Provisional diagnostic criteria for depression of Alzheimer disease: rationale and background. *Am J Geriatr Psychiatry* 2002, 10:129-141.
36. Bohlmeijer E, Roemer M, Cuijpers P, Smit F: The effects of reminiscence on psychological well-being in older adults: A meta-analysis. *Aging Mental Health* 2007, 11:291-300.
37. UVNN, Nieuwboer M, Derksen E: Inventarisatie meetschalen. Nijmegen; 2005.
38. Jongenelis K, Pot AM, Eisses AM, Beekman AT, Kluiter H, Ribbe MW: Prevalence and risk indicators of depression in elderly nursing home patients: the AGED study. *J Affect Disord* 2004, 83:135-142.
39. Kok RM, Nolen WA, Heeren TJ: Venlafaxine versus nortriptyline in the treatment of elderly depressed inpatients. *Int J Geriatr Psychiatry* 2007, 22:1247-1254.

40. Zuidema SU, Derksen E, Verhey FR, Koopmans RT: Prevalence of neuropsychiatric symptoms in a large sample of Dutch nursing home patients with dementia. *Int J Geriatr Psychiatry* 2007, 22:632-638.

41. Evers MM, Samuels SC, Lantz M, Khan K, Brickman AM, Marin DB: The prevalence, diagnosis and treatment of depression in dementia patients in chronic care facilities in the last six months of life. *Int J Geriatr Psychiatry* 2002, 17: 464-472.

42. Verkaik R, Francke AL, van Meijel B, Ribbe MW, Bensing JM: Comorbid depression in dementia on psychogeriatric nursing home wards: which symptoms are prominent? *Am J Geriatr Psychiatry* 2009, 17:565-573.

43. Yesavage JA, Brink TL, Rose T, Lum O, Huang V, Adey M, Leirer V: Development and validation of a geriatric screening scale: a preliminary report. *J Psychiatr Res* 1983, 17:37-49.

44. Smalbrugge M, Jongenelis K, Pot AM, Beekman AT, Eefsting JA: Screening for depression and assessing change in severity of depression Is the Geriatric Depression Scale (30-, 15- and 8-item version) useful for both purposes in nursing home patients? *Aging Mental Health* 2008, 12:244-248.

45. EuroQol Group: EuroQol – a new facility for the measurement of health-related quality of life. Health Policy 1999, 5:21-32.

46. Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, Busschbach van JJ: Kwaliteit van leven meten in economische evaluaties: Het Nederlandse EQ-5D tarief. *Ned Tijdschr Geneeskunde* 2005, 149:1574-1578.

47. Agt van HME, Essink-Bot ML, Krabbe PFM, Bonsel GJ: Test-retest reliability of health state valuations collected with the EuroQol questionnaire. *Social Science & Medicine* 1994, 45: 1537-1544.

48. Coast J, Peters TJ, Richards SH, Gunnell DJ: Use of the EuroQoL among elderly acute care patients. *Qual Life Research* 1998, 7: 1-10.

49. Jönsson L, Andreasen N, Kilander l, Soininen H, Waldemar G, Nygaard H, Winblad B, Jönhagen ME, Hallikainen M, Wimo A: Patient- and proxy-reported utility in Alzheimer disease using the EuroQoL. *Alzheimer Dis Assoc Disord* 2006, 20:49-55.

50. Folstein MF, Folstein SE, McHugh PR: Mini mental state: A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975, 12:189–198.

51. Oostenbrink JB: *Handleiding voor kostenonderzoek: methoden en richtlijnprijzen voor economische evaluaties in de gezondheidszorg.* Diemen: CVZ; 2004.

Results section

Chapter

3

The Nijmegen Observer-Rated Depression scale for detection of depression in nursing home residents

Ruslan Leontjevas, Debby L. Gerritsen, Myrra J.F.J. Vernooij-Dassen, Steven Teerenstra, Martin Smalbrugge, Raymond T.C.M. Koopmans

International Journal of Geriatric Psychiatry, in press

Available online: DOI: 10.1002/gps.2819

Abstract

Objective: To test the accuracy of the Nijmegen Observer-Rated Depression (NORD) scale, a new short scale for screening of depression in nursing home residents with and without dementia

Methods: Cross-sectional study with 103 residents with dementia (N=19 depressed) and 72 residents without dementia (N=10 depressed) was undertaken in 13 Dutch nursing home units. An elderly care physician and a psychologist of each unit assessed residents for the presence of clinical depression. Primary professional caregivers administered the NORD scale.

Results: Five of the six proposed items showed acceptable performance in screening for depression. Receiver operating characteristic analyses revealed significant areas under the empirical curve for the five-item NORD in the total sample (AUC, 0.83; $p < 0.001$), as well as in residents with dementia (AUC=0.84, $p < .001$) and without dementia (AUC, 0.84; $p < 0.001$). The cut-off score of '>1' showed the highest sum of sensitivity (100) and specificity (69) in non-dementia, and '>2' the highest sum of sensitivity (79) and specificity (77) in dementia. The cut-off score of '>1' showed the lowest negative likelihood ratio of 0.0 in non-dementia and of 0.2 in dementia. The highest positive likelihood ratios were found for the cut-off of '>2' in non-dementia (3.4), and for '>4' in dementia (26.5).

Conclusion: The five-item NORD scale showed acceptable accuracy comparable with those of more extensive scales in other studies. It is easy and quickly to administer and can be used for screening of depression in NH residents with or without dementia.

Introduction

Depression is often not recognized and, consequently, undertreated in nursing home (NH) residents.[1-4] To enhance depression recognition in NHs, routine screening procedures with validated instruments have been recommended.[5,6] Indeed, routine screening facilitates recognition and discussion by NH staff of depressive features among residents.[7] However, implementation of the guideline recommendation to screen for depression is challenging in the environment of limited staff time and numerous existing protocols.[8] To facilitate implementation of a screening procedure, it is important to account for possible barriers, such as experienced complexity and staff reluctance.[9]

For depression screening on a regular basis, the use of one instrument for all residents can decrease the complexity of a screening procedure. An observer-rated instrument may be the only option for such a one-scale procedure because many NH residents cannot communicate reliably. Because depressed persons may underreport depression[10] and the nursing staff may be reluctant to use a self-report scale due to expected distress and despair in residents,[11] an observer-rated instrument is a good choice. However, when choosing an observer-rated scale, the following three requirements are important to consider. First, to be used on a regular basis, an instrument should be short and easy to administer. Second, a universal scale for all residents should be based on depressive behavioral symptoms, which do not rely on verbal communication. And, of course, scales accuracy to distinguish depression from non-depression is very important. For this third requirement, it is important to consider that dementia is a highly prevalent condition in NHs and depressive symptoms in persons with dementia can be different from those without dementia.[12,13] Therefore, the screening instrument for NH residents should be accurate in both dementia and non-dementia, or when dementia is not certain.

To the best of our knowledge, to date, a brief observer-rated scale is lacking which is developed for and validated in NH residents and which meets all three requirements. To develop a new scale that meets the requirements, a task-force was appointed by UKON, a Dutch network of ten care organizations and the Department of Primary and Community Medicine of the Radboud University Nijmegen Medical Centre, the Netherlands.

The new Nijmegen Observer-Rated Depression scale (NORD) was developed using the framework of the Brief Observer-Rated Screening Scale for Depression in Elderly Medical Patients (Hammond-scale).[14] The Hammond-scale is short and easy to administer because of its yes/no format to avoid ambiguity inherent in multiple-response. As this scale was constructed for patients with language problems, it consists of observable symptoms that

do not rely on verbal communication. However, the Hammond-scale was neither constructed for, nor validated in NH residents. Moreover, the scale does not account for depression in dementia.[15] To meet the three requirements mentioned above, two of the Hammond-items were incorporated into the NORD scale and four new items were constructed to be used for screening of depression in NH residents with or without dementia.

The aim of our study was twofold: (1) to determine which NORD items can be used in NH residents with or without dementia to screen for depression; and (2) for the NORD consisting of items with acceptable performance, to determine the scale's accuracy, and a range of cut-off scores for screening for depression.

Methods

Participants and design

The sample for the present cross-sectional study was drawn from a larger, ongoing effectiveness study of the multidisciplinary "Act in Case of Depression" care program (AID-study) aimed to identify and treat depression in NH residents [16] The residents were recruited from 7 dementia special care and 6 somatic units participating in the AID study. Inclusion criteria were: written consent and depression assessment that was performed in the AID study prior to intervention for all residents of the 13 cluster randomized units. No additional exclusion criteria were used.

Measures

Nijmegen Observer-Rated Depression (NORD) scale

The Dutch version of the scale was constructed by an expert team of NH professionals consisting of a psycho gerontologist, a psychologist, an elderly care physician, an elderly care nurse and a research assistant. Initially, six items rated as 0=absent, and 1=present were proposed, each of them representing observable symptoms from the diagnostic criteria for major depressive disorder according to the Diagnostic and Statistical Manual, fourth edition, revised (DSM-IV-TR),[17] or from the Provisional Diagnostic Criteria for Depression in Alzheimer Disease (PDCdAD).[15]

Table 3.1
Items of the Nijmegen Observer-Rated Depression (NORD) scale

Instruction: Answer 'yes' when behavior is present, and 'no' when behavior is not present or not applicable to your client. 'Often' means several hours during more days present than not present, during at least the last two weeks.
1. Does the client often look sad, gloomy or cheerless?
2. Does the client often cry or is he/she often emotionally distressed?
3*. Does the client often seem to be restless or irritable (quick-tempered, bad-mood)?
4. Does the client often lack a positive response to social contacts or pleasant events?
5. Does the client often need to be encouraged to do something or participate in joint activities?
6. Has the client often problems with sleep (falling asleep, maintaining sleep, waking up) or appetite (no appetite, unusually hungry)?
*: Item 3 was omitted from the final version

Item construction (Table 3.1): For depressed mood, which is a criterion in both DSM-IV-TR and in PDCdA, two Hammond-scale items were incorporated into the NORD with some textual changes. The term 'miserable' was omitted as not appropriate (item 1), and 'weepy' was replaced by 'emotionally distressed'. The third, new NORD item was developed to assess restlessness and irritability. Restlessness addresses agitation [17] or anxiety, which is common in depression,[18] whereas irritability is the most prevalent symptom in depression in dementia.[19] Item 4 was constructed to include anhedonia of the PDCdAD.[15] The fifth item addresses diminished interest in activities[17] or social isolation and withdrawal.[15] Disruptions in both eating and sleeping are significant predictors for depression[20-22] and they can easily be observed by NH professionals. However, the Hammond sleeping-item did not discriminate between depression and non-depression,[14] whereas sleeping problems are present in only about 25% of depressed residents.[19] Therefore, to increase sensitivity, the NORD team combined sleeping disruptions with disruptions in appetite in one item.

Other measures

Clinical diagnosis of depression served as a gold standard when one of the following diagnoses was established: a major depression according to the DSM-IV-TR;[17] a minor depression with at least two but fewer than five depressive symptoms according to the criteria for major depression, and at least one of the symptoms depressed mood or loss of interest or pleasure; and a depression in dementia according to the Provisional Diagnostic Criteria for Depression in Alzheimer Disease (PDCdAD).[15] The PDCdAD were used

independently of dementia type as it was done in other validation studies,[13,23] or proposed for clinical practice.[5]

Cognitive impairment: The standardized Mini-Mental-State Examination (MMSE)[24] is a scale to assess global cognitive functioning including orientation, registration, attention, calculation, recall and language. On a scale from 0 to 30, a higher score means less cognitive impairment.

Dementia diagnosis was recorded from the medical files. In the Netherlands, dementia in NH residents is mostly assessed in accordance with the DSM-IV-TR[17] by an elderly care physician, or a geriatrician.

Procedure

NH unit managers instructed primary caregivers (credentials are comparable to those of licensed practical nurses) to score the NORD. Doubts about the presence of behaviors should be resolved with other nursing staff members. A group of 70 randomly chosen nursing staff members were asked to estimate how much time it took to fill in the NORD scale. Nursing staff had followed three hours of training in recognition and management of depressive symptoms between two weeks and three months prior to the study. Blinded to the NORD scores, the elderly care physician[25] and psychologist of each unit performed diagnostic procedures guided by the International Diagnostic Checklists for ICD-10 and DSM-IV,[26] which were complemented by the PDCdAD criteria. Depression diagnosis was based on both professionals' collaborative diagnostic activities including the use of chart information, caregiver interview, and examination of the resident (interview, physical examination). A member of the research team, a psychologist or an MSc psychology student, not involved in the daily care at the units, assessed the MMSE.

Ethics

Each participating NH resident and/or the legal representative received written and verbal information prior to the AID study and gave written informed consent. The Medical Ethics Committee (CMO, Arnhem-Nijmegen) rated the study, which was undertaken in accordance with the Helsinki Declaration of 1975, as revised in 1983, and pronounced that it would not be burdensome for the participant.

Statistical analysis

All statistics except those for Receiver Operating Characteristic (ROC) were generated using SPSS 16.0.1 (Chicago, IL).

Descriptive statistics were generated for the whole sample and two subsamples: residents with and without dementia. For the subsamples, depressed and non-depressed residents were compared with respect to age, gender and cognitive impairment. In this regard, and in further analyses, a Chi-Square test was used for categorical data. For other, non-categorical data, which were considered not normally distributed (confirmed by Shapiro-Wilk tests and graph examination), a Mann-Whitney *U* test was used.

Internal Consistency of the NORD was assessed using unstandardized Cronbach's alpha.

NORD items were analysed separately to determine the ratio of depression (sensitivity) and non-depression (1-specificity) with a positive ('yes') item score. The items' accuracy in diagnosing a depression was tested for dementia, non-dementia, and for the total sample using an ROC curves analysis (MedCalc software v.11.1.1.0, Mariakerke, Belgium). The areas under the empirical ROC curve (AUCs) were calculated using trapezoidal rule equivalent to the Wilcoxon estimate based on the method developed by Hanley and McNeil [27]. Items were removed from the scale if they did not meet at least two of the following three criteria: (1) the AUC was larger than area by chance of 0.50, (2) sensitivity and/or specificity was at least 0.70, and (3) item showed no agreement with dementia diagnosis (k-statistic used).

Accuracy of the NORD after removing items that did not show acceptable screening performance was examined using an ROC curves analysis.

Cut-off scores: The range of NORD cut-off scores was examined for dementia, non-dementia, and for the total sample. Sensitivity, specificity, and the positive and negative likelihood ratios (LR+ and LR-) were calculated with the lower and the upper limits of the 95% confidence intervals. The diagnostic odds ratio (LR+ divided by LR-) was determined.[28] A test with an LR+ greater than 10, or an LR- less than 0.1 was considered to have potential to alter clinical decisions, i.e. to, respectively, rule in and rule out depression; and a test with an LR+ between 5 and 10, or an LR-between 0.1 and 0.2 as useful for additional information[28] on screening for depression.

A two sided significance level of 5% was used for all analyses with the exception of item accuracy testing (one-sided).

Results

Of the 242 residents recruited for the current study, 44 (18%) residents died, 16 (7%) moved to another unit, and two (1%) withdrew consent during one to eight months between the consent procedure and the actual assessment in the AID study. Four

residents, all not depressed, were excluded from the study because the NORD was not assessed for administrative reasons. The study population consisted of 53 male (30%) and 122 female residents (Table 3.2). The youngest resident was 45 and the oldest 97 years old. Residents with dementia (N=103) and without dementia (N=72) did not differ in gender (χ^2 [1, N=175]=1.9, *P*=0.161), whereas residents with dementia were older (Mann Whitney [MW] *z*=-2.6, *P*=0.008), and were more cognitively impaired (MW *z*=-8.3, *P*<0.0001).

Of the residents with dementia, nineteen (18.4%) were depressed. Two of them (12%) were male. Depressive residents did not differ from non-depressive in gender (χ^2 [1, N=103]=3.0, *P*=0.085), age (Mann Whitney [MW] *z*=-0.8, *P*=.444), and MMSE score (MW *z*=-0.5, *P*=0.593) (Table 3.2).

Of the 72 persons without dementia, ten (14%) were depressed. Five female residents had major depression according to the DSM-IV-TR and the other four female and one male resident met criteria for a minor depression. Again, depressive residents did not differ from non-depressive residents in gender (χ^2[1, N=72]=3.4, *P*=0.063), and age (MW *z*=-0.3, *P*=0.757), but depressive residents were more cognitively impaired (MW z=-2.3, *P*<0.021) (Table 3.2).

Table 3.2
Demographical and clinical characteristics (N=175)

			Non-dementia			Dementia		
	All residents	Range	Non-depressed n=62	Depressed n=10	Total n=72	Non-depressed n=84	Depressed, n=19	Total, n=103
Male/ Female, *n*	53/122	-	25/37	1/9	26/46	25/59	2/17	27/76
Mean (SD) Age	82.6(8.9)	45–97	79.7(11.5)	80.2(7.9)	79.8(11.0)	84.4(6.8)	86.0(3.7)	84.7(6.3)*
MMSE, n=158	13.5(9.8)	0 – 30	22.5(6.3), n=57	14.2†(10.1), n=9	21.4(7.4)	7.7(7.4), n=75	8.0(5.2), n=17	7.8(7.0)**
NORD-6	2.2(1.9)	0 – 6	1.5(1.8)	3.4(1.1)$^{\#}$	1.8(1.9)	2.1(1.6)	4.2(1.8)$^{\#\#}$	2.4(1.8)**
NORD-5	1.7(1.6)	0 – 5	1.2(1.6)	3.0(1.1)$^{\#}$	1.4(1.6)	1.5(1.3)	3.5(1.5)$^{\#\#}$	1.9(1.5)**

Significance for Mann-Whitney U test: for depression versus non-depression: †*P*<.05; $^{\#}$*P*<.001, $^{\#\#}$*P*<.0001; for dementia versus non-dementia: **P*<.01; ***P*<.0001.
Cr.alpha = unstandardized Cronbach's alpha, whole sample (non-dementia/dementia): NORD-5: 0.74 (0.78/0.70); NORD-6: 0.70 (0.77/0.65)
Abbreviations, MMSE: Mini Mental State Examination; NORD-6: Nijmegen Observer-Rated Depression scale including *irritability* item; NORD-5: final version of the scale.

In total, there were more female than male depressed residents (χ^2 [1, N=175]=6.5, *P*=0.011), and depressive residents did not differ from non-depressive residents in age (MW *z*=-0.4, *P*=0.653) nor MMSE score (MW *z*=-1.8, *P*=0.077).

Caregivers estimated the time they needed to fill in the scale at 2.4 minutes (SD=1.6, range 1.0 to 10.0).

NORD items' figures are summarized in Table 3.3. In the total sample, the first item *sadness* was positively scored (=1) most often in depressed residents (93%) followed by *inactivity* (79%) and *crying* (66%). In non-depressed residents, item 3 *restlessness/irritability* was scored most often (43%). Items 3, 4 (*lack of response*) and 6 (*sleeping and eating problems*) showed non-significant AUCs of 0.55 in this subgroup. However, specificity (81%) for items 4 and 6 was acceptable. Although *restlessness/irritability* in dementia was scored as often as *crying*, and *eating and sleeping problems* (63%) when depression was present, the former item did not show a significant AUC in this group either. Therefore, and due to its low sensitivity and specificity, and the significant agreement with dementia, *restlessness/irritability* was omitted from the scale.

Table 3.3
Performance of the NORD items for dementia (N=103) and non-dementia (N=72)

		Dementia, N=103			Non-dementia, N=72			Total sample, N=175		
Item	k	SENS	1-SP	AUC	SENS	1-SP	AUC	SENS	1-SP	AUC
1 - sadness	.14	89	45	.72†	100	29	.85*	93	38	.77*
2 – crying	-.01	63	17	.73*	70	19	.75†	66	18	.74*
3$^{\#}$ - restlessness/ irritability	.21†	63	52	.55	40	31	.55	55	43	.56
4 – lack of response	.08	53	25	.64**	30	19	.55	45	23	.61$^{\#}$
5 - inactivity	.12	84	43	.71†	70	32	.69**	79	38	.70*
6 – biol.symptoms	.09	63	24	.70†	30	19	.55	52	22	.65†

Abbreviations, k: agreement between dementia diagnosis and NORD item, kappa; SENS: sensitivity; SP: specificity; AUC: area under the empirical ROC curve.
$^{\#}$ Item 3 was omitted from the final scale; Significance: *$*p < .001$, $^{\dagger}P<.01$, $^{**}P<.05$.*

Cronbach's alpha coefficient for the scale with five items (NORD-5 in Table 3.2) was higher in residents without dementia (0.77) than in residents with dementia (0.65).

Scale accuracy: After the *restlessness/irritability* item was omitted, ROC analyses showed moderate discriminative accuracy for the scale in the total sample (AUC=0.83, *SE*=0.05; 95% *CI*: 0.77 to 0.89; *z*=6.9, *P*<0.001), in residents with dementia (AUC=0.84, *SE*=0.06; 95% *CI*: 0.75 to 0.90; *z*=5.7, *P*<0.001), and without dementia (AUC=0.84, *SE*=0.08; 95% *CI*: 0.73 to 0.91; *z*=4.1, P<0.001). Figure 3.1 depicts the ROC curves for the scale in the total sample. Table 3.2 shows higher NORD scores for depressed residents than for non-depressed residents in the total sample (MW: z=-5.8, *P*<0.0001), and in both dementia (MW: z=-4.7, *P*<0.0001) and non-dementia (MW: z=-3.5, *P*<0.001).

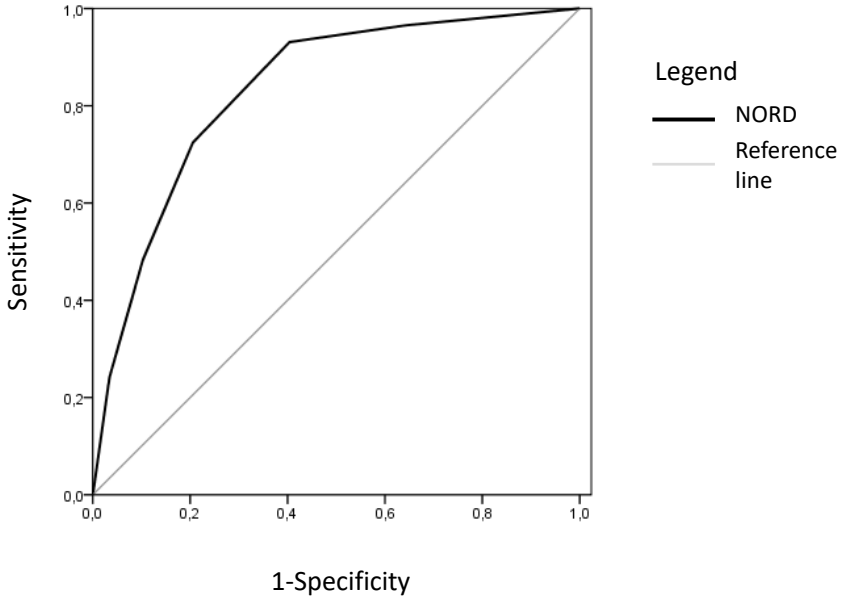


Figure 3.1 Receiver Operating Characteristic Curve for the five-item NORD scale in the heterogeneous sample (N=175) of residents with and without dementia

NORD = Nijmegen Observer-Rated Depression scale; the chance diagonal under the reference line has an area under the curve (AUC) of 0.5. The AUC for the NORD is 0.83 (*SE*= 0.05; 95% *CI*: 0.77 to 0.89; *z*=6.9, *p* <0.001).

Cut-off scores: Table 3.4 presents a range of cut-off scores for the five-item NORD scale. For the total sample, the cut-off score of '>1' showed the highest sum of sensitivity (93) and specificity (60). The score of '>4' yielded the highest LR+ (7.1), and the score of '>0' the lowest LR- (0.09). In non-dementia, the cut-off score of '>1' showed the highest sum of sensitivity (100) and specificity (69) and it had an LR- of 0.0. The highest LR+ was 3.4 for the cut-off '>2'. In dementia, the cut-off score of '>2' had the highest sum of sensitivity

(79) and specificity (77). The score of '>4' had the highest LR+ (26.5), an odds-ratio of 38.4 and an LR- of 0.69. The cut-off scores of '>0' and '>1' had the lowest LR- (0.2).

Table 3.4
Cut-off scores for the five-item NORD scale

	SENS (95% CI)	SP (95% CI)	LR+ (95% CI)	LR- (95% CI)	Odds Ratio
Non-dementia, N=72					
>0	100 (69 - 100)	0 (0 - 6)	1 ()	0 ()	∞
>1*	100 (69 - 100)	69 (56 - 80)	3.3 (2.8 - 3.9)	0 ()	∞
>2	60 (26 - 88)	82 (71 - 91)	3.4 (2.0 - 5.7)	0.5 (0.2 - 1.2)	6.9
>3	30 (7 - 65)	87 (76 - 94)	2.3 (0.9 - 6.0)	0.8 (0.4 - 1.7)	2.9
>4	10 (0.3 - 45)	94 (84 - 98)	1.6 (0.2 - 1.0	1 (0.4 - 2.5)	1.6
Dementia, N=103					
>0	95 (74 - 100)	26 (17 - 37)	1.3 (0.9 - 1.9)	0.2 (0.03 - 1.4)	6.4
>1	89 (67 - 99)	52 (41 - 63)	1.9 (1.5 - 2.4)	0.2 (0.05 - 0.8)	9.4
>2 *	79 (54 - 94)	77 (67 - 86)	3.5 (2.7 - 4.5)	0.3 (0.1 - 0.7)	12.9
>3	58 (34 - 80)	92 (84 - 97)	7 (4.7 - 10.2)	0.5 (0.2 - 1.1)	15.1
>4	32 (13 - 57)	99 (94 - 100)	26.5 (13.7- 51.4)	0.7 (0.10 - 5.0)	38.4
Dementia and non-dementia, N=175					
>0	97 (82 - 100)	36 (28 - 44)	1.5 (1.2 - 1.9)	0.1 (0.01 - 0.7)	15.5
>1*	93 (77 - 99)	60 (51 - 68)	2.3 (2.0 - 2.7)	0.1 (0.03 - 0.4)	19.2
>2	72 (53 - 87)	79 (72 - 86)	3.5 (2.8 - 4.5)	0.4 (0.2 - 0.7)	10.1
>3	48 (29 - 68)	90 (84 - 94)	4.7 (3.2 - 6.9)	0.6 (0.3 - 1.0)	8.1
>4	24 (10 - 44)	97 (92 - 99)	7.1 (3.7 - 13.4)	0.8 (0.3 - 1.9)	8.9

Abbreviations, SENS: sensitivity; SP: specificity; LR+: positive likelihood ratio; LR-: negative likelihood ratio; 95% CI: 95% confidence interval; Odds Ratio = LR+ / LR-.

Discussion

Five of the six proposed NORD items showed acceptable performance in screening for depression in NH residents with and without dementia. The *restlessness/irritability* item did not discriminate between depression and non-depression in dementia or non-dementia. Our results for irritability in residents with dementia seem contradictory to the PDCdAD criteria. Verkaik et al.[19] reported irritability as one of the symptoms most often scored in depression in dementia. However, they did not compare depression with non-depression and they assessed a selected group of residents with high scores on the GDS-15 scale, a scale which may be less useful in residents with severe dementia.[29] Considering that the *restlessness/irritability* item was related to dementia in our study, more research

is needed to determine how the irritability symptom should be assessed for the purpose of depression recognition. In the Hammond et al. study,[14] irritable behavior had low sensitivity and poor inter-rater reliability in non-dementia, but the 'agitation, restlessness and anxiety' item did show acceptable results. Explicitly addressing 'agitation' or 'anxiety' should be considered for future refinement of the NORD scale.

Scale accuracy

After omitting the *restlessness and irritability* item, the ROC analyses revealed that the NORD scale discriminated between depressive and non-depressive residents in both dementia and non-dementia. Using the area under the empirical ROC curve (AUC), which represents a single measure of the discriminative ability of a test, it is possible to compare different scales across the full range of cut-offs.[28] Apparently, the AUC of .83 for the NORD with a 95% confidence interval of 0.77 to 0.89 can be considered to be within the range of AUCs reported for screening instruments based on extensive interviews with patients and/or their proxies. For example, for the 30-item and 15-item Geriatric Depression Scale (GDS), which can be scored by or in an interview with residents,[30] AUCs have been reported ranging from 0.79 to 0.91 and from 0.76 to 0.91 respectively.[31] For the Cornell Scale for Depression in Dementia (CSDD), which is an extensive, observation based instrument for depression in dementia,[32] AUCs of 0.76 and of 0.79 were reported in NH residents with early-onset dementia[23] and in residents with different types of dementia[33,34] respectively. Considering that the discriminative abilities of the NORD in dementia and non-dementia are comparable to those of other more extensive scales, and considering the relatively low labor costs (2-3 minutes to score the scale), the NORD can be used for all residents on a regular basis.

Cut-off scores

For the purpose of screening for depression, we also examined a range of possible cut-off scores for the NORD. Besides choosing a cut-off score with a maximum joint sensitivity and specificity in case of equal costs (gain and losses) of correct and incorrect decisions, likelihood ratios can be used to combine clinical suspicion, i.e. pretest probability, with a test result. For example, the NORD cut-off score of '>1' with a low negative LR can be used by NH professionals to rule out depression, especially when depression is not suspected; the probability of having scores '0' and '1' is low for depressed residents in comparison to non-depressed residents. However, in contrast to non-dementia, zero negative LRs in dementia were not found. Therefore, the NH professional should account for the costs of a false negative result (e.g. suicidal risk) and for the clinical suspicion of depression prior

to testing in dementia. If these costs are very high or depression is highly suspected, additional clinical examination of the clients symptoms may be needed.

To rule in depression, only a cut-off score of '>4' in dementia may have the potential to alter clinical decisions. The high positive LR with pretest probability of 0.5 for this cut-off score indicated almost 27 true positive results for every false positive result. However, the pretest probability of non-depression in residents scoring lower than '5' should be very low because a negative LR for this cut-off score with pretest probability of 0.5 indicated 7 false negative results for every 10 true negative result. For the residents without dementia, the scale did not show a high positive LR, which suggests that, in non-dementia, the scale is less useful for ruling in depression.

In summary, the NORD scores '0' and '1' may be used to rule out depression in dementia or non-dementia. For screening purposes, higher scores may be used depending on the expected costs of false results. Analyses indicated that for equal costs of false negative and false positive results, the optimal cut-off scores of '>1' and '>2' were optimal in non-dementia and dementia, respectively. In dementia, the cut-off score of '>4' can be used to rule in depression.

Limitations

Several aspects of this study should be taken into account when interpreting or generalizing the results. First, the scale was assessed in the Dutch language, and testing of the accuracy in other languages is needed. Second, the training of the nursing staff in detection and management of depressive symptoms may have triggered their awareness and thereby influenced their scoring. However, additional training in scoring or feedback was not provided. Replication with staff who was trained additionally in scoring, or did not have the training at all, as well as comparative validation of the NORD with other scales is needed. The next aspect to be aware of is that we did not restrict our subsamples to a specific disease. We also did not exclude residents with an adjustment disorder, or a depression attributable to a medical condition. On the one hand, false negative numbers are less important than true positive numbers when using a brief instrument followed by a more extensive assessment. On the other hand, to test the universal character of an instrument in a highly heterogeneous NH population, sample restrictions should be avoided. Nevertheless, more research is needed to explore the influence of a specific medical condition or dementia type on the test results.

The next limitation of the study concerns the complexity of depression diagnosis in dementia. Although we always asked two clinicians to diagnose the resident together, it is

not clear whether the gold standard used is reliable in very severe dementia. Lastly, it is not evident whether or not the PDCdAD criteria, which were constructed for Alzheimer's disease, are biased in other types of dementia or comorbidity present in NH residents. To date, these are the only criteria that account for the presence of dementia. Regardless of dementia type, different studies have shown prevalence of depression in dementia comparable to the prevalence in our study.[19,23]

Conclusions

The Nijmegen Observer-Rated Depression (NORD) scale showed intermediate to good discriminative abilities when it was compared to a clinical diagnosis of depression in dementia or non-dementia. Taken together, the results suggest that the five-item NORD scale is quick to administer, shows accuracy comparable with those of more extensive scales, and the scale can be used for screening of depression in NH residents.

Literature

1. Gruber-Baldini AL, Zimmerman S, Boustani M, et al. Characteristics associated with depression in long-term care residents with dementia. *Gerontologist* 2005;45:50-55.
2. Ayalon L, Arean P, Bornfeld H. Correlates of knowledge and beliefs about depression among long-term care staff. *Int J Geriatr Psychiatry* 2008;23:356-363.
3. Snowdon J, Fleming R. Recognising depression in residential facilities: An Australian challenge. *Int J Geriatr Psychiatry* 2008;23:295-300.
4. Kramer D, Allgaier AK, Fejtkova S, et al. Depression in nursing homes: prevalence, recognition, and treatment. *Int J Psychiatry Med* 2009;39:345-358.
5. Koopmans R, Zuidema S, Leontjevas R, et al. Comprehensive assessment of depression and behavioral problems in long-term care. *Int Psychogeriatr* 2010;22:1054-1062.
6. AGS/AAGP. The American Geriatrics Society and American Association for Geriatric Psychiatry recommendations for policies in support of quality mental health care in U.S. nursing homes. *J Am Geriatr Soc* 2003;51:1299-1304.
7. Snowdon J, Rosengren D, Daniel F, et al. Australia's use of the Cornell scale to screen for depression in nursing homes. *Australas J Ageing* 2011;30:33-36.
8. Colon-Emeric CS, Lekan D, Utley-Smith Q, et al. Barriers to and facilitators of clinical practice guideline use in nursing homes. *J Am Geriatr Soc* 2007;55:1404-1409.
9. Grol RP, Wensing M. *Implementatie: Effectieve verbetering van de patiëntenzorg.* Amsterdam: Elsevier gezondheidszorg, 2010.
10. Snow AL, Kunik ME, Molinari VA, et al. Accuracy of self-reported depression in persons with dementia. *J Am Geriatr Soc* 2005;53:389-396.
11. Hammond MF. Doctors' and nurses' observations on the Geriatric Depression Rating Scale. *Age Ageing* 2004;33:189-192.
12. Olin JT, Katz IR, Meyers BS, et al. Provisional diagnostic criteria for depression of Alzheimer disease: rationale and background. *Am J Geriatr Psychiatry* 2002;10:129-141.

13. Barca ML, Selbaek G, Laks J, et al. The pattern of depressive symptoms and factor analysis of the Cornell Scale among patients in Norwegian nursing homes. *Int J Geriatr Psychiatry* 2008;23:1058-1065.
14. Hammond MF, O'Keeffe ST, Barer DH. Development and validation of a brief observer-rated screening scale for depression in elderly medical patients. *Age Ageing* 2000;29:511-515.
15. Olin JT, Schneider LS, Katz IR, et al. Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002;10:125-128.
16. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011;11:91.
17. APA. Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association, 2000.
18. Prado-Jean A, Couratier P, Druet Cabanac M, et al. Specific psychological and behavioral symptoms of depression in patients with dementia. *Int J Geriatr Psychiatry* 2010;25:1065-1072.
19. Verkaik R, Francke AL, van Meijel B, et al. Comorbid depression in dementia on psychogeriatric nursing home wards: which symptoms are prominent? *Am J Geriatr Psychiatry* 2009;17:565-573.
20. Djernes JK. Prevalence and predictors of depression in populations of elderly: a review. *Acta Psychiatr Scand* 2006;113:372-387.
21. Cole MG, Dendukuri N. Risk factors for depression among elderly community subjects: a systematic review and meta-analysis. *Am J Psychiatry* 2003;160:1147-1156.
22. Martin L, Poss JW, Hirdes JP, et al. Predictors of a new depression diagnosis among older adults admitted to complex continuing care: implications for the depression rating scale (DRS). *Age Ageing* 2008;37:51-56.
23. Leontjevas R, van Hooren S, Mulders A. The Montgomery-Asberg Depression Rating Scale and the Cornell Scale for Depression in Dementia: a validation study with patients exhibiting early-onset dementia. *Am J Geriatr Psychiatry* 2009;17:56-64.
24. Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry* 1991;148:102-105.
25. Koopmans RTCM, Lavrijsen J, Hoek JF, et al. Dutch Elderly Care Physician: A new generation of nursing home physician specialists. *J Am Geriatr Soc* 2010;58:1807-1809.
26. Hiller W, Zaudig M, Mombour W. *IDCL: International diagnostic checklists for ICD-10 and DSM-IV*. Seattle: Hogrefe & Huber Publishers, 1996.
27. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
28. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003;29:1043-1051.
29. Debruyne H, Van Buggenhout M, Le Bastard N, et al. Is the geriatric depression scale a reliable screening tool for depressive symptoms in elderly patients with cognitive impairment? *Int J Geriatr Psychiatry* 2009;24:556-562.
30. Yesavage JA, Brinka TL, Roseb TL, et al. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* 1983;17:37-49.
31. Jongenelis K, Pot AM, Eisses AM, et al. Diagnostic accuracy of the original 30-item and shortened versions of the Geriatric Depression Scale in nursing home patients. *Int J Geriatr Psychiatry* 2005;20:1067-1074.
32. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23:271-284.
33. Barca ML, Engedal K, Selbaek G. A reliability and validity study of the cornell scale among elderly inpatients, using various clinical criteria. *Dement Geriatr Cogn Disord* 2010;29:438-447.

34. Leontjevas R, Gerritsen DL, Vernooij-Dassen M, et al. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents in dementia. *Am J Geriatr Psychiatry* In press.

Chapter 4

A comparative validation of the proxy-based Montgomery-Åsberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents with dementia

Ruslan Leontjevas, Debby L. Gerritsen, Myrra J.F.J. Vernooij-Dassen, Martin Smalbrugge, Raymond T.C.M. Koopmans

American Journal of Geriatric Psychiatry, in press

Available online: DOI: 10.1097/JGP.0b013e318233152b

Abstract

Objective: (1) to compare the accuracy of the Montgomery-Åsberg Depression Rating Scale (MADRS) and the Cornell Scale for Depression in Dementia (CSDD) in nursing home residents with dementia when professional caregivers are the only available source of information, and (2) to explore different methods to account for missing items.

Design: cross-sectional design.

Setting: nursing homes.

Participants: 101 residents with dementia.

Measurements: NH residents with dementia were assessed on the presence of clinical depression using Provisional Diagnostic Criteria for Depression of Alzheimer's Disease. The MADRS and CSDD were administered in a structured interview with professional primary caregivers.

Results: Receiver operating characteristic analyses revealed no significant differences between areas under the empirical curve for MADRS and CSDD. Imputation of a lowest possible item score for missing items revealed larger areas than three other methods (significant result only for CSDD). A MADRS cut-off score of '>13' yielded the highest sum of sensitivity (78%) and specificity (66%). A CSDD cut-off score of '>6' yielded the highest sum of sensitivity (94%) and specificity (49%). Both scales showed high negative predictive values up to 100% and low positive predictive values not exceeding 50%.

Conclusion: The proxy-based MADRS and CSDD did not differ in distinguishing depressed from non-depressed NH residents and may be used for screening purposes. For missing items, imputation of a lowest possible item score may be applied. The MADRS and CSDD may be better used for ruling out rather than for ruling in depression.

Introduction

Although depression in dementia is treatable,[1,2] numerous studies have shown that depression is often not recognized and is undertreated in nursing home (NH) residents with dementia.[3,4] This is worrying because depression in residents with dementia is very common, with rates up to 86%,[5] has a significant impact on the quality of life,[6-8] and is associated with other poor outcomes such as accelerated cognitive decline[9] and mortality.[10,11] Cognitive decline and language problems complicate recognition of depression in dementia:[12] The need for information from proxies' observations is apparent. However, proxies may also have difficulty in evaluating certain symptoms in residents with severe communicative problems, such as presence of suicidal or pessimistic thoughts.[13] Therefore, considering that the use of screening instruments enhances depression recognition in dementia,[12] NH professionals need valid proxy-based instruments with practical procedures to account for possible missing items.

The Cornell Scale for Depression in Dementia (CSDD),[14] which employs information from proxies, is the most often used scale in clinical trials on depression in long-term care settings[15] and is widely applied in clinical practice. Considering that some CSDD items may be highly sensitive to other comorbid conditions in NH residents, the use of depression scales with less somatic/vegetative items has been suggested for NH residents with dementia.[16,17] However, validation studies that compare the performance of such scales with the CSDD are very scarce, especially research that uses the Provisional Diagnostic Criteria (PDC) for Depression of Alzheimer's Disease[18] as a gold standard. Depressive symptoms in persons with dementia can be different from those without dementia,[17,19] and the PDC account for that, whereas the other often-used standardized criteria for depression[20,21] do not.

Recently, the Montgomery- Åsberg Depression Rating Scale (MADRS), with less emphasis on somatic symptoms, was compared with the CSDD in NH residents exhibiting early onset dementia ([EOD] dementia onset before age of 65 years),[13] using the PDC as gold standard. In contrast to the CSDD, the MADRS was not originally designed to specifically assess depression in dementia. The MADRS is widely used in the clinical context because of its sensitivity to change following treatment, whereas the CSDD's sensitivity to change is limited.[22] In the EOD study, solely primary professional caregivers were interviewed to score the scales. Interestingly, the MADRS performed better than the CSDD on concurrent validity against the PDC.

To the best of our knowledge, there is only one more validation study[23] in NH residents against the PDC. Barca et al.[23] examined the CSDD and reported cut-off scores against the PDC for a mixed sample of NH residents and hospital patients of geriatric psychiatry. According to the original scale instruction,[24] interviews of both residents and caregivers were used for scoring. More validation research against the PDC using only one source of information, namely proxies, will provide NH professionals with cut-off scores that can be used for all residents, independent of the severity of cognitive and/or communicative problems. Furthermore, comparing the performance of the proxy-based MADRS and CSDD in the general NH population with dementia may guide NH professionals in the choice of a screening instrument.

As mentioned, NH professionals also need a procedure to account for symptom-evaluation difficulties by proxies resulting in missing items. Many comprehensive algorithms for imputation of missing items exist[25] but most are complex, and therefore less applicable by NH professionals in daily practice. A conservative imputation with a lowest possible item score produced acceptable results in the EOD study[13] but, to date, this method has not been compared with other easy-to-perform methods. Therefore, the aim of our study was twofold: (1) to compare test accuracy of the proxy-based MADRS and CSDD, and (2) to compare several easy-to-perform imputation methods for missing items.

Methods

Participants and Design

The sample for the present cross-sectional study was drawn from a larger, ongoing effectiveness study of the multidisciplinary care program, "Act in Case of Depression" (AID-study), aimed to identify and treat depression in NH residents.[26] All data reflect baseline and preintervention assessments. The residents were recruited from seven of the 16 cluster-randomized dementia special care units participating in the AID study. Inclusion criteria were: written consent by the legal representative on behalf of the client and presence of dementia as stated in the personal file available on the ward.

Measures

The Cornell Scale for Depression in Dementia (CSDD)[14,27] consists of 19 items, each rated as 0=absent, 1=mild or intermitted, and 2=severe. A higher total scale score indicates more severe depressive symptoms.

The Montgomery-Åsberg Depression Rating Scale (MADRS)[28,29] consists of 10 items, each rated on seven intensity grades from absent=0 to severe=6. The items'

intensity grades 0, 2, 4, and 6 are formulated separately with behavioral examples for each item. A higher total score indicates more severe depressive symptoms.

Clinical diagnosis of depression in dementia: An expert group initiated by the American National Institute of Mental Health proposed the Provisional Diagnostic Criteria (PDC) for Depression of Alzheimer's Disease.[18] The PDC are derived from the Diagnostic and Statistical Manual of Mental Disorders (DSM)[20] but they are broader and more sensitive.[23, 30] At least three symptoms are required during the same 2-week period and must include either depressed mood or anhedonia. In contrast to the DSM criteria, the symptoms of irritability and social isolation or withdrawal are added and concentration problems are omitted. Other symptoms are comparable to DSM criteria: disruption in appetite, sleeping problems, psychomotor changes, fatigue or loss of energy, feelings of worthlessness, hopelessness, or excessive or inappropriate guilt, and recurrent thoughts of death, suicidal ideation, plan or attempt. In our study and in accordance with other studies,[23, 31, 32] criterion B was considered as "All criteria are met for dementia" instead of "… for dementia of Alzheimer type".

Cognitive impairment: The standardized Mini-Mental-State Examination (MMSE)[33] is a widely-used continuous measure of global cognitive functioning including orientation, registration, attention, calculation, recall and language. On the scale from 0 to 30, a higher score means less cognitive impairment.

Dementia diagnosis was recorded from the unit medical files. In the Netherlands, dementia in NH is generally assessed in accordance with the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition Text revision (DSM-IV-TR)[20] by an elderly care physician, or geriatrician.

Procedure

The elderly care physician and psychologist of each unit performed diagnostic procedures guided by the International Diagnostic Checklists for ICD-10 and DSM-IV[34] that were adapted for depression in dementia according to the PDC. Depression diagnosis was based on their collaborative diagnostic activities including the use of chart information, caregiver interview, and examination of the resident (interview, physical examination). The MMSE was administered in a structured interview with the resident, and the MADRS and CSDD with the primary professional caregiver (credentials can be compared with those of licensed practical nurse in United States). All interviewed caregivers were also asked to report on their opinions about the MADRS and CSDD. About half of the residents (51%) were randomly evaluated with the CSDD first, whereas the other half (49%) were

administered the MADRS first. All interviews were conducted by a psychologist or by graduate psychologists in their final year of M.Sc., none of whom were involved in the intervention program. The interviewers had been trained in administering the scales.

Ethics

Each legal representative received written and verbal information prior to the AID study and gave informed consent in writing. When communication was possible, residents received a short explanation before they were interviewed. All residents had the opportunity to decline the interview. The Medical Ethics Committee (CMO, Arnhem-Nijmegen) rated the study and declared it to be not burdensome for the participant.

Statistical Analysis

All statistics, except those for receiver operating characteristic (ROC), were generated using SPSS 16.0.1 (Chicago, IL).

Descriptive statistics were generated. Depressed and non-depressed residents were compared on age, gender, whether or not residents had at least one missing item on the MADRS or CSDD, and whether or not an interview was possible. Interview was considered possible when an MMSE score was at least higher than '0'. In this regard, and in further analyses, a Chi-Square (χ^2) test was used for categorical data. For other, non-categorical data which were considered not normally distributed (confirmed by Shapiro-Wilk test and graph examination), a Mann-Whitney U test was used.

Scale indices: For the MADRS and CSDD, easy-to-perform imputation methods were run resulting in four indices per scale: (1) 'lowest score' (LS) indices with missing items filled up with '0', i.e. the lowest possible item scores; (2) 'middle-most score' (MS) indices with the score of '3' for MADRS and '1' for CSDD missing items; (3) 'highest score' (HS) indices with the score of '6' for MADRS and '2' for CSDD missing items, and (4) 'extrapolated score' (ES) indices with missing items filled in with an average score of the remaining items.[25]

Internal consistency of the scales was assessed using unstandardized Cronbach's alpha.

Congruent validity was assessed by the Spearman's correlation (r_s) between the MADRS and CSDD. A correlation less than 0.30 was regarded as weak, between 0.30 and 0.69 as moderate and of 0.70 and higher as strong.[35]

Concurrent (criterion) validity was assessed by ROC curves analyses using MedCalc software v.11.1.1.0 (Mariakerke, Belgium). The areas under the empirical ROC curve (AUCs, using trapezoidal rule equivalent to the Wilcoxon estimate) were calculated on the basis of the method developed by Henley and McNeil.[36] A test with an AUC greater than

0.9, between 0.7 and 0.9, or between 0.5 and 0.7 was considered as having respectively high, moderate, or low discriminative accuracy.[37] The significance of the AUCs for each MADRS and CSDD index, and the significance of the differences between two AUCs was calculated with the z-test.[38] AUCs for all scale indices were computed and the scale index with a larger AUC was considered the better measure to distinguish depressed from non-depressed residents.

The cut-off score with the highest sum of sensitivity and specificity was considered the optimal score to distinguish between depressed and non-depressed residents, assuming equally-weighted false positive and false negative errors.[39] For the scale indices with the largest AUCs, a range of cut-off scores is presented. Positive and negative predictive values (PPV and NPV), which are important parameters for clinical practice, were calculated using sample prevalence. For screening purposes, avoiding false negatives may be important; therefore, cut-off scores with a high sensitivity and NPV were determined. For diagnostic purposes, a cut-off score should maximize the specificity and PPV because avoiding misdiagnosis is important.

Results

Of the 117 residents recruited for the current study, sixteen residents died or moved to another unit before they could be assessed. The study sample consisted of 28 male (27.7%) and 73 female residents (Table 1). Of the 101 residents from seven care units (8 to 22 residents per unit), 18 (17.8%) were diagnosed as depressed. Dementia type, as specified in personal file, was as follows (first diagnosis N [n depressed]): Alzheimer's disease N=25[6], Vascular dementia N=24[1], other type N=7[2], and unknown N=35[9]. Depressed and non-depressed residents did not differ significantly in age and gender (Table 1), and they did not differ in whether they could be interviewed (N=47, MMSE>0) or could not (χ^2 [1, N=101]=0.72, *P*=0.397). When an interview with a resident was possible, depressed (n=10) and non-depressed (n=37) residents did not differ in the MMSE score (Table 4.1).

Table 4.1
Demographical and clinical characteristics (N=101)

	Non-depressed *n*=83	Depressed *n*=18	Mann Whitney U test, z	*P* (2-tailed)	All residents	Range
Male/Female, *n*	26/57	2/16	-	.082[a]	28/73	-
Age, M (SD)	83.4 (6.8)	85.7 (3.7)	-1.0	.336[b]	83.8 (6.4)	64.8 – 94.2

MMSE>0, n=47	11.1 (5.9), n=37	10.5 (4.5), n=10	- .5	.611[b]	11.2 (5.6)	2 - 23
MADRS *M(SD)*						
- LS	12.5 (8.4)	18.7 (7.2)	-3.1	.002[b]	13.6 (8.5)	0 – 35
- MS	13.5 (9.1)	19.0 (7.6)	-2.8	.006[b]	14.5 (9.0)	0 - 38
- HS	14.6 (10.6)	19.3 (8.0)	-2.5	.012[b]	15.4 (10.3)	0 – 48
- ES	13.2 (9.2)	18.9 (7.8)	-2.8	.005[b]	14.2 (9.2)	0 – 39
CSDD *M(SD)*						
- LS	7.4 (4.4)	11.8 (3.7)	-3.8	<.001[b]	8.2 (4.6)	0 – 20
- MS	8.2 (5.0)	12.1 (3.7)	-3.2	.001[b]	8.9 (5.0)	0 – 21
- HS	8.9 (5.9)	12.3 (3.7)	-2.8	.005[b]	9.5 (5.7)	0 – 26
- ES	7.8 (4.8)	11.9 (3.8)	-3.5	<.001[b]	8.6 (4.9)	0 – 21

Significance[a] : mentioned for χ^2 (1, N=101) = 3.0 ; significance [b]: mentioned for Mann-Whitney test
Abbreviations, MMSE > 0: residents who could be interviewed, i.e. a score higher than '0' on the Mini-Mental-State Examination; MADRS: Montgomery-Åsberg Depression Rating Scale; CSDD: Cornell Scale for Depression in Dementia;
Indices accounting for missing items, LS: 'lowest possible item score'; MS: 'middle-most score'; HS: 'highest score'; ES: 'extrapolated score'.

Missing Items. Professional caregivers could not assess at least one symptom in 36 residents (36%). Seventeen residents (17%) missed an average of 1.8 of the MADRS items (SD=1.5; range = 1 to 6), and 28 residents missed an average of 2.4 of the CSDD items (SD=2.2; range 1 to 10). Depressed and non-depressed residents did not differ in having a missing item or not (χ^2 [1, N=101]=0.10, *P*=0.751). The number of missing items was related to the score based on the other present items for the CDSS (N=101, r_s=0.32, *P*=0.001) but not for the MADRS (N=101, r_s= 0.08, *P*=0.411).

Internal Consistency. Cronbach's alpha coefficients for the MADRS indices varied from 0.74 to 0.79 and for the CSDD from 0.70 to 0.77 (Table 4.2). The coefficients did not increase by more than 0.05 when any of individual items were deleted.

Table 4.2
Internal consistency and results of ROC analyses for the MADRS and CSDD indices (N=101)

	Cr. alpha (Unst.)	AUC (*SE*)	95% CI	z	*P*	Cut-off	SENS %	SP %	PPV %	NPV %
MADRS										
- LS	.74	.73 (.07)	.63 - .82	3.2	.001	>13	78	66	33	93
- MS	.78	.71 (.07)	.61 -.80	2.8	.004	>13	83	61	32	94
- HS	.79	.69 (.07)	.59 - .78	2.6	.010	>13	83	58	30	94
- ES	.79	.71 (.07)	.61 - .80	2.9	.004	>13	78	64	32	93
CSDD										
- LS	.70	.79 (.07)	.69 - .86	4.3	<.001	>6	94	49	29	98
- MS	.74	.74 (.07)	.65 - .83	3.4	<.001	>5	100	41	27	100
- HS	.77	.71 (.07)	.62 - .80	2.9	.003	>8	89	53	29	96
- ES	.74	.76 (.07)	.67 - .84	3.8	<.001	>6	94	48	28	98

Significance: P-value for z-test; Cut-off: optimal cut-off score with the maximum of sensitivity and specificity. *Abbreviations*, MADRS: Montgomery- Åsberg Depression Rating Scale; CSDD: Cornell Scale for Depression in Dementia; Cr. Alpha (Unst.): Cronbach alpha unstandardized; AUC (SE): area under the curve (Standard Error); CI: confidence interval; SENS: sensitivity; SP: specificity; PPV, NPV: negative and positive predictive values; Indices accounting for missing items, LS: 'lowest possible item score'; MS: 'middle-most score'; HS: 'highest score'; ES: 'extrapolated score'

Congruent validity

The correlations between MADRS and CSDD indices were strong and significant (N=101; Presented pairwise, for LS: r_s=0.75; MS: r_s=0.73; HS: r_s=0.70; ES: r_s =0.75, *P*<0.0001 for all pairs).

Concurrent Validity

The AUCs for all scale indices were significantly larger than AUC=0.50, i.e. area by chance (Table 2).

MADRS Indices: Depressed residents scored higher on the mean scores of all MADRS indices (one-tailed; Table 4.1). The AUCs for the MADRS indices (Table 4.2) varied from 0.69 (HS) to 0.73 (LS). Pairwise comparison did not reveal any significant differences between these AUCs. The cut-off score of '>13' showed the highest sum of sensitivity (78% to 83%) and specificity (58% to 66%) for all MADRS indices.

CSDD Indices: Depressed residents scored higher on the mean scores of all CSDD indices (one-tailed; Table 4.1). The AUC for the LS index (Table 2) was significantly larger than that for the MS (ΔAUC=0.04, z=2.2, *P*=0.027), the HS (ΔAUC=0.07, z=2.5, *P*=0.011), and the ES (ΔAUC=0.02, z=2.1, *P*=0.032). The optimal CSDD cut-off scores varied from '>5' (MS) to '>8' (HS). Almost all depressed residents (sensitivity of 89% to 100%) and only about half of the non-depressed residents (specificity of 41% to 53%) were identified correctly.

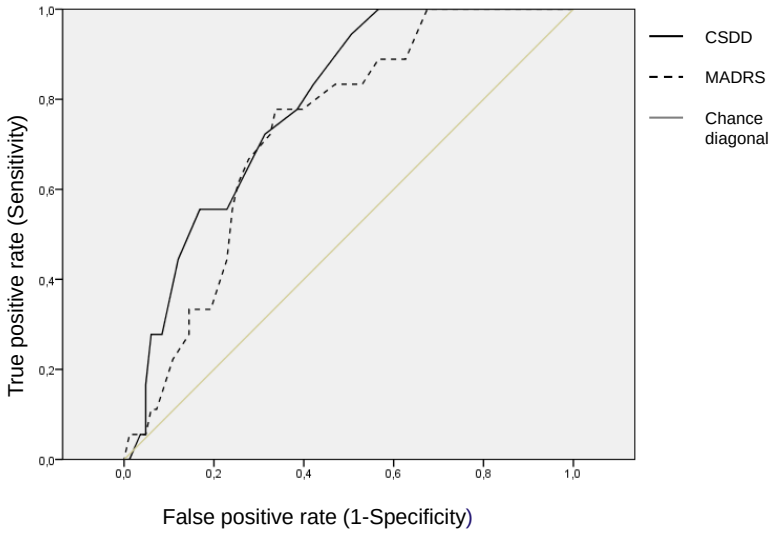


Figure 4.1 Receiver Operating Characteristic Curve for the LS indices of the Montgomery-Åsberg Depression Rating Scale (MADRS) and Cornell Scale for Depression in Dementia (CSDD)

LS-index: missing items imputed with a lowest possible item score. The chance diagonal has an area under the curve (AUC) of 0.5. The AUC for the MADRS is 0.73 and AUC for the CSDD is 0.79. The difference between MADRS and CSDD is not significant.

MADRS versus CSDD: Figure 4.1 depicts the ROC curves for the LS indices. The AUC difference between the MADRS and CSDD was not significant (ΔAUC=0.05, z=0.8, *P*=0.407), nor were pairwise differences between other scales' indices (MS, HS, ES) significant.

Cut-off scores: For the LS indices, a range of cut-off scores is presented in Tables 4.3 and 4.4. Both scales showed high NPV's up to 100% and low PPV's up to 50%. A MADRS cut-off score of '>7' and a CSDD score of '>5' had the highest sum of sensitivity and NPV. The MADRS cut-off score of '>34', and CSDD of '>14' (not shown) had the highest sum of PPV (both scales: 50%) and specificity (MADRS: 99%, CSDD: 94%).

Table 4.3
Cut-off scores for the Montgomery- Åsberg Depression Rating Scale (using LS index)

	>6	>7	>8	...	>12	>13 *	>14	>15
Sensitivity	100	100	89		78	78	72	67
Specificity	27	33	37		60	66	67	72
PPV	23	24	24		30	33	33	34
NPV	100	100	94		93	93	92	91

PPV= positive predictive value; NPV = negative predictive value; * - maximum of sensitivity and specificity; LS-index: missing items imputed with a lowest possible item score.

Table 4.4
Cut-off scores for the Cornell Scale for Depression in Dementia (using LS index)

	>4	>5	>6 *	>7	>8	>9	>10
Sensitivity	100	100	94	83	78	72	56
Specificity	30	43	49	58	61	69	77
PPV	24	28	29	30	30	33	35
NPV	100	100	98	94	93	92	89

PPV= positive predictive value; NPV = negative predictive value; * - maximum of sensitivity and specificity; LS-index: missing items imputed with a lowest possible item score.

Discussion

With regard to the first aim of this study, both MADRS and CSDD performed moderately in distinguishing depressed from non-depressed NH residents with dementia when only one source of information was available, namely that of professional caregivers. The ROC analyses revealed that the scales did not significantly differ in their performance against the diagnostic criteria for depression in dementia. Although concurrent validity was acceptable, both the MADRS and CSDD had a combination of a high NPV and a low PPV. That suggests that the MADRS and CSDD may be better used for ruling out rather than for ruling in depression. For the LS indices, the MADRS scores of '7' and lower and the CSDD scores of '5' and lower can be used to rule out depression without missing depressed residents. Higher scores should encourage professionals to further examine the presence of depression and may be used for screening purposes. However, the higher scores can

not be used for diagnostic purposes. This implies that using cut-off scores for starting a treatment, let alone pharmacological treatment that can show side effects,[40] would be inappropriate without further mental health evaluation.

Ideally, professionals should always account for the costs (gains and losses) of correct and incorrect decisions, costs that may differ in different contexts. If the costs of a false positive error and a false negative error are approximately equal, a cut-off score that maximizes the proportion of correct test results may be seen as optimal. The MADRS cut-off score yielding the highest discriminative ability was '>13'. It showed a sensitivity of 78% and a specificity of 66%. The 'screening' score of '>14' proposed in the EOD study[13] showed a comparable specificity of 67% but a higher sensitivity of 92%.The optimal CSDD cut-off score was '>6', which is equal to the score with the highest discriminative ability obtained by Barca et al. for patients with Alzheimer's disease.[23] The '>6' score detected a large part of the depressed residents (sensitivity of 94%), but only half of the non-depressed residents (specificity of 49%). Considering that both scales did not significantly differ on their AUCs, the scales correlated strongly with each other, and the scoring time is mostly the same, the CSDD optimal cut-off score with higher sensitivity may be preferred to avoid the high costs of a false negative error, whereas the MADRS cut-off score may be preferred to avoid the high costs of a false positive error. Yet, future research is needed to determine optimal cut-off scores from different perspectives, such as patients, society and health service providers.[39]

Interestingly, in the present study, the AUC of 0.79 for the CSDD was about the same as in the EOD study[13] (0.76), whereas for the MADRS, the AUC was smaller in the present study (0.73 versus 0.87). We noticed that in daily practice the proxy-based CSDD was the most frequently used instrument on dementia special care units, and fourteen of the 27 interviewed caregivers in the present study found it easier to score the CSDD than MADRS. Only seven caregivers appreciated the MADRS and its behavioral examples, whereas most caregivers in the EOD study preferred the MADRS to the CSDD and described the former as 'easier to understand' and 'to link to a person'.[13] More research is needed to examine whether setting differences and staff characteristics, or the time of dementia onset and other residents' characteristics may explain the discrepancy for the MADRS performance between the studies.

Regarding our second aim, the finding that in more than one third of the NH residents (36%) caregivers could not evaluate at least one of the depressive symptoms, underscores the importance of exploring different ways to account for missing items. To our knowledge, this is the first validation study of depression scales in NH residents with

dementia to include an evaluation of different methods to account for missing items. The value of such research is apparent for NH professionals, as different ways to account for missing items may lead to different total scores that imply the necessity to use different optimal cut-off scores. For the CSDD, a conservative method of assigning the missing items with the lowest possible item score (LS) showed a significantly better performance on a ROC analysis than other imputation methods. Interestingly, the AUC for the LS index was the same as in the study of Barca et al.,[23] in which the score was based on information from interviews of both residents and their proxies. This agreement suggests that solely using information from caregivers and conservatively filling in the missing items with a 'zero' – because of lack of ability to evaluate them - does not necessarily result in a less valid measurement. Although not significant, the LS method also performed better for the MADRS, suggesting that this method may also be used for this scale.

It is important to mention that solely using information from caregivers may reduce resource utilization in NH practice. However, an interview with a resident may be necessary to better understand the resident's depressive symptoms and, whenever an interview is possible, we do not recommend skipping it. More extensive research is needed to determine whether, and to what extent, the lack of a resident's interview influences the scales' psychometrical qualities for distinguishing depressed from non-depressed NH residents with dementia. Furthermore, for a proxy-based instrument, the acceptable number of missing items should be determined, and imputation techniques that may be used for other purposes than clinical, such as research or epidemiological, should be examined.

Several aspects of this study should be taken into account when interpreting the results. First, because of practical problems, two psychologists involved in the assessment of depression according to the PDC had access to the scores of the MADRS (N=11) or CSDD (N=9) prior to the diagnostic procedure. However, hierarchical logistic regression models with the MADRS and CSDD scores as predictors did not reveal that the non-blinded procedure was related to depression diagnosis (data not shown). A second consideration is that we did not restrict our sample to a specific type of dementia, as the NH population is heterogeneous in dementia type, and the exact course of the condition is often unknown. Interestingly, in the study of Barca et al.,[23] the cut-off scores obtained in a sample of patients with Alzheimer's disease differed from those obtained for participants with different types of dementia. Although the AUCs for the CSDD obtained by Barca et al. were virtually identical to our AUC, more research with proxy-based instruments may be needed to obtain cut-off scores for specific dementia types. Third, considering the small

numbers of residents per unit, we did not control for possible site effects on validity measures. Additional analyses (not shown) did not reveal significant differences between units in depression rate (χ^2 test), and number of missing items or mean MADRS and CSDD scores (analysis of variance test). And last but not the least, it is not evident whether or not the PDC that were constructed for Alzheimer's disease, are biased in other types of dementia present in NH residents. To date, these are the only criteria that account for the presence of dementia, and different studies show,[13,31] regardless of the type of dementia, that the prevalence of PDC depression was comparable to the prevalence in our study. Nevertheless, the PDC validation for different types of dementia is recommended.

Conclusions

The fact that interviewing the resident was not possible in at least half of our sample underlines the need for the validation of proxy-based depression scales. The results of our study showed that the CSDD and MADRS may be used for screening purposes in NH residents with dementia when only one source of information is available, namely professional caregivers. In addition, it was demonstrated that different imputation methods for missing items might lead to different optimal cut-off scores which may have implications for NH care. There was no significant difference between the MADRS and CSDD in distinguishing depressed from non-depressed NH residents, and both scales may be applied in residents who cannot be interviewed. Using these scales with the provided cut-off scores may enhance depression recognition in daily NH practice.

Literature

1. Lyketsos CG, Olin J: Depression in Alzheimer's disease: overview and treatment. *Biol Psychiatry* 2002; 52:243-252
2. O'Connor DW, Ames D, Gardner B, et al: Psychosocial treatments of psychological symptoms in dementia: a systematic review of reports meeting quality standards. *Int Psychogeriatr* 2009; 21:241-251
3. Gruber-Baldini AL, Zimmerman S, Boustani M, et al: Characteristics associated with depression in long-term care residents with dementia. Gerontologist 2005; 45 Spec No 1:50-55
4. Snowdon J, Fleming R: Recognising depression in residential facilities: An Australian challenge. *Int J Geriatr Psychiatry* 2008; 23:295-300
5. Wright SL,Persad C: Distinguishing between depression and dementia in older persons: Neuropsychological and neuropathological correlates. *J Geriatr Psychiatry Neurol* 2007; 20:189-198
6. Wetzels RB, Zuidema SU, de Jonghe JF, et al: Determinants of quality of life in nursing home residents with dementia. *Dement Geriatr Cogn Disord* 2010; 29:189-197

7. Gonzalez-Salvador T, Lyketsos CG, Baker A, et al: Quality of life in dementia patients in long-term care. *Int J Geriatr Psychiatry* 2000; 15:181-189
8. Smalbrugge M, Pot AM, Jongenelis L, et al: The impact of depression and anxiety on well being, disability and use of health care services in nursing home patients. *Int J Geriatr Psychiatry* 2006; 21:325-332
9. Rapp MA, Schnaider-Beeri M, Wysocki M, et al: Cognitive decline in patients with dementia as a function of depression. *Am J Geriatr Psychiatry* 2010; PDF Only. doi:10.1097/JGP.1090b1013e3181e1898d1090
10. Suh GH, Kil Yeon B, Shah A, et al: Mortality in Alzheimer's disease: a comparative prospective Korean study in the community and nursing homes. *Int J Geriatr Psychiatry* 2005; 20:26-34
11. Barca ML, Engedal K, Laks J, et al: A 12 months follow-up study of depression among nursing-home patients in Norway. J Affect Disord 2010; 120:141-148
12. Teresi J, Abrams R, Holmes D, et al: Prevalence of depression and depression recognition in nursing homes. *Soc Psychiatry Psychiatr Epidemiol* 2001; 36:613-620
13. Leontjevas R, van Hooren S, Mulders A: The Montgomery-Asberg Depression Rating Scale and the Cornell Scale for Depression in Dementia: a validation study with patients exhibiting early-onset dementia. *Am J Geriatr Psychiatry* 2009; 17:56-64
14. Alexopoulos GS, Abrams RC, Young RC, et al: Cornell Scale for Depression in Dementia. *Biol Psychiatry* 1988; 23:271-284
15. Hyer L: Depression in Long Term Care. *Clin Psychol - Sci Pr* 2005; 12:280-299
16. Kurlowicz LH, Evans LK, Strumpf NE, et al: A psychometric evaluation of the Cornell Scale for Depression in Dementia in a frail, nursing home population. *Am J Geriatr Psychiatry* 2002; 10:600-608
17. Olin JT, Katz IR, Meyers BS, et al: Provisional diagnostic criteria for depression of Alzheimer disease: rationale and background. *Am J Geriatr Psychiatry* 2002; 10:129-141
18. Olin JT, Schneider LS, Katz IR, et al: Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002; 10:125-128
19. Barca ML, Selbaek G, Laks J, et al: The pattern of depressive symptoms and factor analysis of the Cornell Scale among patients in Norwegian nursing homes. *Int J Geriatr Psychiatry* 2008; 23:1058-1065
20. Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision, Washington, DC, American Psychiatric Association, 2000
21. International Statistical Classification of Diseases, 10th Revision (ICD-10) Geneva, Switzerland, World Health Organization, 1992
22. Colasanti V, Marianetti M, Micacchi F, et al: Tests for the evaluation of depression in the elderly: A systematic review. *Arch Gerontol Geriatr* 2010; 50:227-230
23. Barca ML, Engedal K, Selbaek G: A reliability and validity study of the cornell scale among elderly inpatients, using various clinical criteria. *Dement Geriatr Cogn Disord* 2010; 29:438-447
24. Alexopoulos GS: The Cornell Scale for Depression in Dementia: Administration & scoring guidelines. December 2, 2002. Available at: *http://scalesandmeasures.net.* Accessed June 10, 2010.
25. Schafer J, Graham J: Missing data: Our view of the state of the art. *Psychological methods* 2002; 7:147-177
26. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011; 11. doi:10.1186/1471-244X-11-91.
27. Dröes RM. The Dutch version of the Cornell scale for depression in dementia. In: Dröes R, editor. *Amsterdamse Ontmoetingscentra, een nieuwe vorm van ondersteuning voor dementerende mensen en hun verzorgers.* Nijkerk: Thesis Publishers; 1996, p.157.

28. Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979; 134:382-389
29. Hartong EGTM GJ: Montgomery Åsberg Depression Rating Scale. [De Montgomery Åsberg beoordelingsschaal voor depressie]. *Tijdschr Psychiatrie* 1985 27:657-669
30. Vilalta-Franch J, Garre-Olmo J, Lopez-Pousa S, et al: Comparison of different clinical diagnostic criteria for depression in Alzheimer disease. *Am J Geriatr Psychiatry* 2006; 14:589-597
31. Verkaik R, Francke AL, van Meijel B, et al: Comorbid depression in dementia on psychogeriatric nursing home wards: which symptoms are prominent? *Am J Geriatr Psychiatry* 2009; 17:565-573
32. Koopmans R, Zuidema S, Leontjevas R, et al: Comprehensive assessment of depression and behavioral problems in long-term care. *Intern Psychoger* 2010; 22:1054-1062
33. Molloy DW, Alemayehu E, Roberts R: Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry* 1991; 148:102-105
34. Hiller W, Zaudig M,Mombour W: IDCL: International diagnostic checklists for ICD-10 and DSM-IV, Seattle, Hogrefe & Huber Publishers, 1996
35. Jackson S: Research methods and statistics: A critical thinking approach. Belmont CA, Wadsworth Cengage Learning, 2008
36. Hanley JA, McNeil BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36
37. Fischer JE, Bachmann LM, Jaeschke R: A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003; 29:1043-1051
38. Hanley JA, McNeil BJ: A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148:839-843
39. Smits N, Smit F, Cuijpers P, et al: Using decision theory to derive optimal cut-off scores of screening instruments: an illustration explicating costs and benefits of mental health screening. *Int J Method Psych* 2007; 16:219-229
40. Mark TL, Joish VN, Hay JW, et al: Antidepressant Use in Geriatric Populations: The Burden of Side Effects and Interactions and Their Impact on Adherence and Costs. *Am J Geriatr Psychiatry* 2011; 19:211

Chapter

5

A comparative validation of the abbreviated Apathy Evaluation Scale (AES-10) with the Neuropsychiatric Inventory Apathy subscale against diagnostic criteria of apathy

Ruslan Leontjevas, Alexandra Evers-Stephan, Martin Smalbrugge, Anne Margriet Pot, Viviane Thewissen, Debby L. Gerritsen, Raymond T.C.M. Koopmans

Journal of American Medical Directors Association, 2011, 13(3):308.e1-6

Abstract

Objective*:* To compare the Neuropsychiatric Inventory apathy subscale (NPIa) with the abbreviated Apathy Evaluation Scale (AES-10) on discriminant validity and on their performance to distinguish residents as apathetic or non-apathetic.

Design: Cross-sectional design.

Setting: Nursing homes.

Participants: 100 residents of 4 dementia special care units (n=58) and 3 somatic units.

Measurements: Primary professional caregivers were interviewed to score the AES-10 and NPIa. The elderly care physician and the psychologist of each unit examined residents for clinical apathy using diagnostic criteria.

Results*:* The AES-10 and NPIa correlated moderately with each other (r_S=0.62, P<0.0001). The AES-10 correlated weakly (r_S=0.27, P=0.024) and NPIa moderately (r_S=0.46, P=0.001) with the Cornell Scale for Depression in Dementia. Receiver operating characteristic analysis showed an area under the curve (AUC) of 0.72 (P<0.01) for AES-10, and 0.67 (P<0.05) for NPIa. The AES-10 produced higher sums of sensitivity and negative predictive value than the NPIa. Explorative analyses revealed that both instruments produced higher scores in dementia independently of having an apathy diagnosis, whereas AUCs were significant in non-dementia (AES-10: AUC=0.88, P<0.001; NPIa: AUC=0.77, P=0.023), but not in dementia.

Conclusion*:* Both the AES-10 and NPIa may be used to distinguish apathetic from non-apathetic residents in a heterogeneous sample with and without dementia, or in residents without dementia. The AES-10 may be preferable to the NPIa apathy subscale when ruling out or screening for apathy. The performance of the scales against diagnostic criteria of apathy in dementia need to be further examined.

Introduction

Apathy is increasingly recognized as a clinically significant neurobehavioral phenomenon.[1,2] According to an extensive review, apathetic features are very common in numerous neurological and psychiatric conditions, with the highest prevalence of 84.1% for apathy found in nursing home (NH) residents.[3] Apathy is associated with a number of adverse outcomes, such as high caregiver burden,[4] poor treatment response, reliance on caregivers to initiate activities of daily living, low quality of life, and more rapid cognitive and functional decline.[3,5] Considering the high prevalence rate of apathy in NH residents, the numerous adverse effects, and the potential treatment opportunities through psychological and pharmacological interventions,[6-8] NH professionals need reliable and valid instruments to assess apathy. Recognizing apathy in clinical practice is important because apathy can be mistaken for loss of ability to perform activities, or be mislabelled as depression. At the same time, treatment of depression may induce apathy.[9] Hence, it is important to differentiate apathy from depression.[1,4] Because NH residents have high comorbidity and are heterogenic in physical illnesses, and in cognitive and functional impairments, apathy instruments with robust psychometric measures across different medical conditions are needed.

Assessment of apathy has always been complicated because a common definition of apathy was lacking and there was no consensus as to which diagnostic criteria or screening instruments should be used.[10,11] Fortunately, in 2009, an international task force including European, American and Australian experts introduced *Diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders (DCA)*. The DCA have not yet been adopted in existing nomenclature, but the criteria are the first and only for a formal diagnosis of apathy to be agreed on by an international task force (including members of the Association Française de Psychiatrie Biologique, the European Psychiatric Association, the European Alzheimer's Disease Consortium and experts from Europe, Australia and North America).[11] The use of the criteria can facilitate communication in research and practice, and contribute to reliable case description and identification. This can have therapeutic consequences resulting in quality of life improvement. The DCA's validity has been established for several neuropsychiatric disorders, such as Alzheimer's disease, mixed dementia, mild cognitive impairment, Parkinson disease, schizophrenia and major depression.[2,12] Consequently, several instruments that have been developed to assess apathy can now be validated against this "gold" standard.

In a recent review, Clarke et al.[10] analyzed fifteen apathy scales and reported that both the apathy subscale of the Neuropsychiatric Inventory (NPIa) and the Apathy Evaluation Scale (AES) revealed the most robust psychometric measures across different medical conditions. The NPI[13,14] is an informant-based rating scale developed to measure neuropsychiatric symptoms, including apathy (NPIa), in patients who suffer from Alzheimer's disease and other dementia syndromes. It can additionally be used in clients with various brain injuries and other diseases. Concurrent validity of the NPIa against the DCA has been examined and was satisfactory.[2,12] The scale has often been used both in clinical practice as a screening instrument, and in research to determine the convergent validity of other scales. A nursing home version of the NPI (NPI-NH)[14,15] is available. The Apathy Evaluation Scale (AES-18) was developed to characterize and quantify apathy in individuals aged 55 and older.[16] Lueken and colleagues[17] refined the AES scale for the NH population by eliminating 8 items that either had lost specificity owing to the mainly externally driven context in NHs, or were difficult to measure in residents with severe cognitive deficits. The abbreviated version, AES-10, performed well on different psychometric properties and correlated as strongly with the NPIa as the original AES-18. Lueken et al.[17] concluded that the easily and quickly completed, user-friendly AES-10 can be used in the NH population to assess apathy.

To our knowledge, both the AES-10 and NPIa have not yet been validated against the DCA in NH residents and, consequently, researchers and NH professionals are missing data to determine which instrument – AES-10 or NPIa – is better able to distinguish apathetic from non-apathetic residents. The first objective of this study was to compare the discriminant validity of the NPIa and AES-10 using the Cornell Scale of Depression in Dementia so as to examine the capacity of the scales to distinguish apathy from depression. The second aim was to evaluate the performance of the AES-10 and NPIa in distinguishing residents in DCA-apathetic and non-apathetic. Our third aim was to determine cut-off scores against the diagnostic criteria for apathy.

Methods

Participants, Design, and Ethics

This cross-sectional study was carried out with NH residents who had been participating in a larger, ongoing effectiveness study aimed at identifying and treating depression in NHs.[18] Participants from seven randomly chosen units were examined for the presence of apathy according to the DCA. All data reflect baseline and preintervention observations. Inclusion criteria were the following: (1) written informed consent given by the NH-resident him/herself, or by his/her legal guardian, because of the physical or psychological constraints of the resident, and (2) an independent ("blind") assessment of the participant according to the DCA. The Medical Ethics Committee (CMO, Arnhem-Nijmegen) rated the study and decided that as it is a non-invasive intervention study, it is not burdensome for the participant.

Measures

Abbreviated Apathy Evaluation Scale (AES-10): The abbreviated scale, AES-10, by Lueken and colleagues[17] consists of 10 items. Each item of the observational scale gives an example of apathetic behavior. The answer categories are the following: 1 = not at all characteristic, 2 = slightly characteristic, 3 = somewhat characteristic, and 4 = a lot characteristic, resulting in a scale ranging from 10-40. A higher total AES-10 score indicates more apathetic behavior.

Neuropsychiatric Inventory (nursing home), (NPI-NH): Validity and reliability of the Dutch version of the observational NPI has previously been established.[19] The *Apathy/Indifference* subscale (NPIa), consisting of a general screening question and 8 subquestions rated by the informant, a nursing staff member, with 'yes' or 'no', was used in this study. The frequency and severity of apathy is determined when subquestions confirm the screening item. The NPIa score is obtained by multiplying severity (rated as 1-3) by frequency (rated as 1-4) of the apathetic behaviour. The lowest possible score is 0 and the highest possible score is 12; a higher score represents more severe apathy.

Diagnostic Criteria for Apathy in Alzheimer's Disease and Other Neuropsychiatric Disorders (DCA). Apathy is diagnosed when (1) the patient exhibits lack of motivation; (2) there is at least one symptom in at least two of the following three domains, for a period of at least four weeks, and which is present most of the time: loss of, or diminished goal-directed behavior, and/or goal-directed cognitive activity, and/or emotion; (3) the symptoms cause clinically significant impairment in personal, social,

occupational or other important areas of functioning; and (4) the symptoms are not explained by or due to physical disabilities, motor disabilities, diminished level of consciousness or physiological effects of a substance. The criteria are assembled in a way that allows them to be applied not only in the context of Alzheimer's disease or other neuropsychiatric disorders, but also non-dementia related disorders.[11]

Cornell Scale for Depression in Dementia (CSDD): The observational CSDD[20,21] consists of 19 items, each rated as 0=absent, 1=mild or intermittent, and 2=severe. A higher total score indicates more severe depressive symptoms than a lower score. The scale has been validated in patients both with dementia[20] and without dementia[22] as well as in NH residents.[23,24] In our study, unstandandardized (standardized) Cronbach's alpha of the CSDD scale was 0.77 (0.77).

Procedure

The AES-10, NPIa and CSDD were administered during a structured interview with primary professional caregivers. The interviews were conducted by six MSc psychology students, one research assistant and one psychologist, all of whom had been trained in administering the scales. For half of the NH residents, the NPIa was administered first and the AES-10 was administered second. For the other half, vice versa. The elderly care physician (N=7) and the psychologist (N=7) of each unit, who were both involved in the daily care, performed the diagnostic procedure collaboratively. They were guided by a DCA-based template while they were blinded for the results of the NPIa and the AES-10. Their diagnostic activities included the use of chart information, caregiver interview, and examination of the resident (interview, physical examination). Information about the diagnosis of dementia and age were obtained from the client medical file.

Statistical Analyses

Statistical analyses were performed using SPSS 16.0 (Chicago, IL). MedCalc 13.1 (Mariakerke, Belgium) was used for receiver operating characteristic analysis and correlations with 95% confidence intervals (95% CI).

Descriptive statistics were generated for apathetic and non-apathetic residents and compared on gender, age, diagnosis of dementia, mean scores of the AES-10 and CSDD, and mode of the NPIa. An independent t-test was used for normally distributed data, a Mann-Whitney *U* test for not normally distributed data (confirmed by Shapiro-Wilk test and graph examination), and the chi-square test for categorical data.

Internal consistency of the multi-item scale AES-10 was assessed by Cronbach's alpha.[25] Because the NPIa was not designed as a multi-item scale, internal consistency was not calculated for this instrument.

Discriminant validity: A Spearman coefficient (r_s) for both scales with the CSDD were calculated in order to examine discriminant validity. A correlation of less than 0.30 was regarded as weak, between 0.30 and 0.69 as moderate and of 0.70 and higher as strong.[26] Weak correlations were thought to reflect good discriminant validity. The difference between two r_s was considered significant when the 95% CI did not overlap.

Congruent validity: A Spearman coefficient (r_s) was calculated for AES-10 and NPIa in order to assess congruent validity.

Concurrent (criterion) validity and cut-off scores: Receiver operating characteristic (ROC) curves with 95% CIs were calculated to estimate sensitivity and specificity of the NPIa and AES-10 against the DCA at the various cut-off scores. The cut-off score with the highest sum of sensitivity and specificity was considered as the optimal cut-off to distinguish NH residents as apathetic or non-apathetic assuming that negative effects of a false positive error and a false negative error are approximately equal. Positive and negative predictive values (PPV and NPV) were calculated using the sample prevalence of DCA apathy.

Areas under the empirical ROC curve (AUCs) were calculated using trapezoidal rule equivalent to the Wilcoxon estimate, based on the method developed by Hanley and McNeil.[27] An AUC of 0.7 or higher means that the test performs fairly well to excellent.[28] The significance of the difference between the AUC for the AES-10 and the AUC for the NPIa was calculated with the z-test.[29] The scale with the largest AUC was considered the better test to distinguish residents with apathy from residents without apathy.

Explorative analyses: Apathetic symptoms have been associated with dementia and age.[5,11,16,30] Dementia and age were, as well as gender, examined as possible correlates of the AES-10 and NPIa. Explorative ROC analyses were applied, accounting for revealed correlates.

Results

In total, informed consent was obtained from 109 (67%) of the 163 residents of the participating units. Six of the 109 residents died and two moved to non-participating units prior to assessment. One resident was not assessed for apathy due to administrative errors.

Descriptive statistics and clinical characteristics of the 100 included participants are shown in Table 5.1. Male residents were slightly younger (M=82, SD=7.4) than female residents (M=85, SD=6.5) (t(98)=-1.9, *P*=0.049). In total, 14 residents (14 %) were diagnosed with apathy according to the DCA. Apathetic clients were younger than non-apathetic clients (t(98)=3.2; *P*=0.002) and had higher scores on the AES-10 (MW z=2.59, *P*=0.009) and NPIa (MW z=2.46, *P*=0.014). The differences between apathetic and non-apathetic clients were not significant regarding gender (χ^2 [1, N=100]=0.569, *P*=0.451), the CSDD score (data available for 13 apathetic and 58 non-apathetic residents, MW z=-1.93, *P*=.053), and having a diagnosis of dementia (χ^2 [1, N=100]=0.005, p = 0.944). Six of the apathetic residents (43%) were evaluated as not having apathy according to the NPIa (score = 0).

Congruent validity was moderate. The AES-10 and NPIa correlated significantly with each other (r_s =0.62, N=100, *P*<0.0001; 95% CI: 0.48 to 0.73).

Discriminant validity did not differ significantly between the AES-10 and NPIa. Both instruments correlated significantly with the CSDD, and the 95% CI of the two Spearman rho's overlapped (AES-10: r_s =0.27, N=71, *P*=0.024; 95% CI: 0.04 to 0.47; NPIa: r_s =0.46, N=71, *P*<0.001; 95% CI: 0.26 to 0.63).

Internal consistency was high for the AES-10. Unstandardized (standardized) Cronbach's alpha of the AES-10 scale was 0.95 (0.95). The coefficients did not increase by more than 0.05 if any of the items were deleted. The mean inter-item correlation of the AES-10 was 0.632.

Table 5.1
Demographics and Characteristics of Nursing Home Residents (N=100)

	Total	Apathetic clients (N=14)	Non-apathetic clients (N=86)	Clients with dementia (N=58)	Clients without dementia, (N=42)
Age, mean (SD), [range]	84.0 (6.9) [65-97]	78.8(7.7) [67-92]	84.9(6.4) [65-97]	84.2(7.5) [65-97]	83.7(6.1) [69-94]
Sex, male (%) / female	34(34)/66	6(42)/8	28(32)/59		
Dementia diagnosis					
- No dementia (%)	42(42)	6(43)	36(42)	-	42
- Dementia (%)	58 (58)	8(100)	50(100)	58(100)	-
Alzheimer's disease (%[1])	17	4(50)	13(26)	17(29)	-
Vascular dementia (%[1])	11	-	11 (22)	11(20)	-
Others (%[1])	4	1(13)	3(6)	4(7)	-
Mixed (%[1])	8	2(25)	6(12)	8(14)	-
Not specified (%[1])	18	1(13)	17(34)	18(31)	-
AES-10, mean (SD), [range]	25.2(9.4) [10-40]	31.1(7.2) [17-40]	24.2(9.3) [10-40]	29.5(7.5) [14-40]	19.3(8.4) [10-40]
NPIa, mode, [range]	10 [1-25]	14 [0-12]	0 [0-12]	0 [0-12]	0 [0-12]
CSDD, mean (SD), [range], n	9.1(5.5) [1-25], 71	12.2(6.4) [4-25], 13	8.5(5.1) [1-20], 58	8.5(4.9) [1-20], 54	10.9(7.0) [1-25], 17

[1]Percentage of dementia; Dementia diagnosis, Not specified: dementia without further description or no specific diagnosis made; Others: Korsakov syndrome, Lewy body dementia, Parkinson Disease Dementia, Frontotemporal dementia and Huntington's Chorea;
Abbreviations, AES-10: Apathy Evaluation Scale; NPIa: Neuropsychiatric Inventory apathy subscale; CSDD: Cornell Scale for Depression in Dementia

Concurrent (criterion) validity

The ROC curve analysis is depicted in figure 5.1. The AUC for AES-10 was 0.72 and significantly greater than .50, i.e. area by chance (SE= .08; 95% CI: 0.62 to 0.80; z=2.66, *P*<0.01). The AUC for NPIa was .67 (SE= 0.08; 95% CI: 0.57 to 0.76; z=2.00, *P*<0.05). The difference between both AUCs was 0.05 and not significant (ΔAUC =0.05, SE= 0.08; 95% CI: -0.11 to 0.21; z=0.59, *P*=0.559).

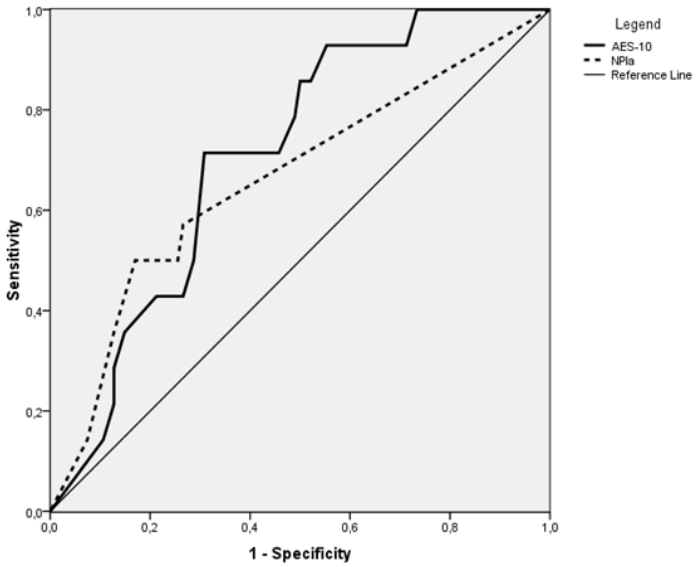


Figure 5.1 Receiver Operating Characteristic Curve for the AES-10 and NPIa in the heterogeneous sample (n = 100) of residents with and without dementia.

The chance diagonal under the reference line has an AUC of 0.5. The AUC for the AES-10 is 0.72 and the AUC for the NPIa is 0.67. AES, Apathy Evaluation Scale; NPIa, Neuropsychiatric Inventory apathy subscale.

Cut-off scores

The cut-off scores for both scales are represented in Tables 5.2 and 5.3. The optimal cut off score of '>29' for the AES-10 had the highest sum of sensitivity (0.71) and specificity (0.70) meaning that a score of 29 or less indicates the absence of apathy and a score of 30 or higher is indicative for apathy. The score had the highest PPV (0.28). The optimal NPIa cut-off score of '>4' showed the highest sum of sensitivity (0.50) and specificity (0.84).

Table 5.2
Cut-off Scores for the AES-10

Cut-off score	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Total sample, N=100				
>28	71 (42 - 92)	69 (58 - 78)	27 (14 - 44)	94 (84 - 98)
>29 *	71 (42 - 92)	70 (59 - 79)	28 (14 - 45)	94 (85 - 98)
>30	50 (23 - 70)	72 (61- 81)	23 (10 - 41)	90 (80 - 96)
Dementia, n=58				
>28	88 (47 - 100)	52 (37 - 66)	23 (10 - 41)	96 (81 - 100)
>29 *	88 (47 - 100)	54 (39 - 68)	23 (10 - 42)	96 (82 - 100)
>30	63 (25 - 92)	56 (41 - 70)	19 (6 - 38)	90 (74 - 98)
Non-dementia, n=42				
>20	100 (54 - 100)	67 (49 - 81)	33 (13 - 59)	100 (86 - 100)
>21 *	100 (54 - 100)	75 (58 - 88)	40 (16 - 68)	100 (87 - 100)
>22	83 (36 - 100)	81 (64 - 92)	42 (15 - 72)	97 (83 - 100)

PPV/NPV: positive/negative predictive value; * maximum of sensitivity and specificity;
95% CI: 95% confidence interval.

Table 5.3
Cut-off scores for the NPI Apathy Subscale

Cut-off score	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Total sample, N=100				
>3	50 (23 - 77)	79 (69 - 87)	28 (12 - 49)	91 (82 - 96)
>4 *	50 (23 - 77)	84 (74 - 91)	33 (15 - 57)	91 (83 - 96)
>6	36 (13 - 65)	88 (80 - 94)	33 (12 - 62)	89 (81 - 95)
Dementia, n=58				
>3	50 (16 - 84)	68 (53 - 81)	20 (6 - 44)	90 (75 - 97)
>4 *	51 (16 - 84)	76 (62 - 87)	25 (7 - 52)	91 (77 - 97)
>6	38 (9 - 76)	84 (71 - 93)	27 (6 - 61)	89 (77 - 97)
Non-dementia, n=42				
>0*	67 (22 - 96)	89 (74 - 97)	50 (16 - 84)	94 (80 - 99)
>1	50 (12 - 88)	89 (74 - 97)	43 (10 – 82)	91 (77 - 97)
>2	50 (12 - 88)	92 (78 - 98)	50 (12 - 88)	92 (77 - 98)

PPV/NPV: positive/negative predictive value; * maximum of sensitivity and specificity;
95% CI: 95% confidence interval.

Correlates (explorative analyses, results only shown for significant values)

The AES-10 and NPIa were not related to age and gender. In residents without apathy (N=86), the instruments showed higher scores in dementia than in non-dementia (AES-10: MW z=5.62, *P*<0.0001; NPIa: MW z=2.58, *P*=0.010). In residents with apathy (N=14), dementia was not found to be associated with the AES-10 and NPIa scores, which may be because of the small subsample size.

When residents were split into two groups, those without dementia (N=42, Table 5.1), and those with dementia (N=58), they did not differ in gender, age and CSDD scores. However, residents with dementia scored higher on the AES-10 (MW z=5.42, *P*<0.0001) and NPIa (MW z=2.08, *P*=0.037). Furthermore, in residents without dementia, residents with apathy showed higher scores on the AES-10 (MW z=2.96, *P*=0.003) and NPIa (MW z=3.04; *P*=0.002) when compared to non-apathetic residents. On the contrary, when dementia was present, the scores of AES-10 and NPIa did not differentiate apathetic and non-apathetic residents.

Considering that dementia, independent from apathy, was associated with the AES-10 and NPIa scores, ROC analyses were performed separately for residents with and without dementia. In residents without dementia the AUCs for both instruments were significant (AES-10: AUC=0.88; z=4.05, *P*=0.0001; NPIa: AUC=0.77, z=2.26, *P*=0.023), whereas the difference between both AUCs was not significant. The optimal cut off score of '>21' for the AES-10 scale had the highest sum of sensitivity (1.00) and specificity (0.75), and also the highest NPV (1.00). The optimal NPIa cut-off score of '>0' showed the highest sum of sensitivity (0.67) and specificity (0.89). In residents with dementia, the AUCs for both instruments did not reach significance.

Discussion

With regard to our first aim, discriminant validity was good for the AES-10, as the scale correlated weakly with the CSDD depression scale. On the contrary, the NPIa correlated moderately with the CSDD, suggesting that the NPIa is more related to depressive symptoms than the AES-10. However, the advantage of the AES-10 above the NPIa in terms of discriminant validity was not significant, as the 95% confidence intervals for the correlation coefficients of both scales with the CSDD overlapped.

Concurrent validity against the diagnostic criteria for apathy was satisfactory for both the AES-10 and the NPIa. The instruments performed well in distinguishing between apathetic and non-apathetic residents, and apathetic residents scored higher on the AES-10 and NPIa than non-apathetic residents. ROC analysis showed that the AUC for the AES-10 was

significant and even exceeded 0.700, which may be considered fairly well.[28] Although the AUC for the NPIa was significant, it did not reach 0.700. It should be noted that the difference between the both AUCs was not significant, which indicates that the scales may not differ in their performance to distinguish apathetic and non-apathetic residents. A comparison of the instruments showed that they correlated moderately suggesting that the scales are not fully congruent but measure about the same.

In line with the results of the study by Lueken et al.,[17] internal consistency of the AES-10 was strong. A Cronbach's alpha above 0.9 suggests that the items strongly interrelate and some of them may even be unnecessary. This was also confirmed by the mean inter-item correlation (0.63) which was above the ideal range of correlations (0.2 to 0.4).[31]

NH residents with dementia and without dementia did not differ on prevalence of apathy, gender, age and depression scores. However, the residents with dementia in both the whole sample and in the non-apathetic subgroup had higher scores on the NPIa and AES-10 than residents without dementia. The higher scores on apathy instruments in dementia are in line with other research.5, 11, 16, 30 Nevertheless, it is quite remarkable that apathetic and non-apathetic residents differed on AES-10 and NPIa when dementia was not present, but did not differ in scores on both scales when dementia was present. According to the DCA, apathy must be excluded when a diminished level of consciousness, which may accompany dementia,32 explains the apathetic symptoms. The AES-10 and NPIa do not have this exclusion criterion. This, as well as other exclusion criteria of the DCA such as physical and motor disabilities, may explain why the DCA-apathy prevalence in this study was relatively low compared with the range of numbers from 13% up to 90% reported in studies using screening scales3. Additional research with larger groups for dementia and non-dementia is needed to determine whether a diminished level of consciousness, or physical and motor disabilities could explain our finding that the AES-10 and NPIa did not discriminate apathetic residents when dementia was present.

With regard to the obtained cut-off scores, it should be noted that both the AES-10 and NPIa showed a combination of a high NPV and a low PPV, suggesting that the instruments may be better used for ruling out than for ruling in apathy. Considering that the AES-10 produced higher sums of sensitivity and NPV than the NPIa, the use of the AES-10 may be preferable for screening purposes when recognizing as many apathetic residents as possible is important. An AES-10 score of '16' or lower may be used to rule out apathy without missing apathetic residents. A score higher than '21' should encourage clinicians to further examine the presence of apathy. Assuming equal costs (gains and losses) of a false positive error and a false negative error, the cut-off score of '>21' may be considered

as optimal for residents without dementia. The score showed the highest sum of sensitivity (100%) and specificity (75%), maximizing the proportion of correct test results. In the total sample of residents with and without dementia, an optimal AES-10 cut-off score of '>29' correctly identified 72% of the NH-residents with apathy and 70% without apathy. In contrast, the performance of the NPIa was less acceptable, as only 67% of the NH-residents with apathy were identified correctly in non-dementia using optimal NPIa cut-off score of '>0', and 50% were identified correctly in the total sample using the cut-of score of '>4'. Nevertheless, these cut-off score showed acceptable specificity varying from 76 to 89%, which may be because of the NPIa eight control subquestions that follow the positive screening question.

Screening for apathy in NH residents is important because apathy is associated with a number of adverse outcomes and its recognition may be helpful in differential diagnosis and formulation of treatment strategies.[11] To the best of our knowledge, this is the first comparative validation study of screening instruments against diagnostic criteria for apathy. There are several aspects that should be considered when interpreting our results. First, given the small size of the groups with and without dementia in our study, the ROC analyses for dementia and non-dementia should be interpreted as exploratory, and more validation research is needed with larger samples, especially in NH residents with dementia. Second, the study was performed using the scales in Dutch. Testing of the accuracy in other languages is needed. Third, we used a heterogeneous sample with regard to medical conditions because restricting the sample would not reflect the NH population which is very diverse in terms of physical illness, functional impairment and type of dementia. However, additional validation for specific medical conditions may be advised. Another limitation of our study may concern the use of the DCA as gold standard. The criteria have not been validated for the NH population, but for several neuropsychiatric and neurobehavioral disorders in outpatients and hospital patients.[2,12] Indeed, more validation research of the criteria and the scales in the NH setting is certainly needed to underpin our conclusions about applicability of the AES-10 and NPIa in the general NH population. Furthermore, we did not control for the prevalence artificially and reported the proportion of subjects with correct diagnoses (PPV and NPV). However, it is important to mention that prevalence can influence predictive values. To avoid sample selection bias, participants in this study were recruited from randomly selected units, but selection may have played a role because informed consent was obtained from two thirds of the residents of the units. More extensive epidemiological research is needed to

determine whether the DCA-prevalence reported in this study can be generalized to the NH population.

Conclusions

Both AES-10 and NPIa distinguished apathetic from non-apathetic NH residents in a heterogeneous sample of residents with and without dementia, and in a subsample without dementia. The AES-10 may be preferable to the NPIa apathy subscale when ruling out or screening for apathy. The applicability of the scales in dementia need to be further examined in larger samples.

Literature

1. Starkstein SE, Leentjens AF. The nosological position of apathy in clinical practice. *J Neurol Neurosurg Psychiatry* 2008; 79:1088-1092.
2. Mulin E, Leone E, Dujardin K, et al. Diagnostic criteria for apathy in clinical practice. *Int J Geriatr Psychiatry* 2011; 26:158-165.
3. Van Reekum R, Stuss DT, Ostrander L. Apathy: Why care? *J Neuropsychiatry Clin Neurosci* 2005; 17:7-19.
4. Landes AM, Sperry SD, Strauss ME, et al. Apathy in Alzheimer's disease. *J Am Geriatr Soc* 2001; 49:1700-1707.
5. Tagariello P, Girardi P, Amore M. Depression and apathy in dementia: Same syndrome or different constructs? A critical review. *Arch Gerontol Geriatr* 2009; 49:246-249.
6. Hsieh C-J, Chang C, Su S-F, et al. Reminiscence Group Therapy on Depression and Apathy in Nursing Home Residents With Mild-to-moderate Dementia. *J Exp Clin Med* 2010; 2:72-78.
7. Drijgers RL, Aalten P, Winogrodzka A, et al. Pharmacological treatment of apathy in neurodegenerative diseases: a systematic review. *Dement Geriatr Cogn Disord* 2009; 28:13-22.
8. Roth RM, Flashman LA, McAllister TW. Apathy and its treatment. *Curr Treat Options Neurol* 2007; 9:363-370.
9. Barnhart WJ, Makela EH, Latocha MJ. SSRI-induced apathy syndrome: A clinical review. *J Psychiatr Pract* 2004; 10:196-199.
10. Clarke DE, Ko JY, Kuhl EA, et al. Are the available apathy measures reliable and valid? A review of the psychometric evidence. *J Psychosom Res* 2011; 70:73-97.
11. Robert P, Onyike CU, Leentjens AF, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatry* 2009; 24:98-104.
12. Drijgers RL, Dujardin K, Reijnders JS, et al. Validation of diagnostic criteria for apathy in Parkinson's disease. *Parkinsonism Relat Disord* 2010; 16:656-660.
13. Cummings JL, Mega M, Gray K, et al. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology* 1994; 44:2308-2314.
14. Wood S, Cummings JL, Hsu MA, et al. The use of the Neuropsychiatric Inventory in nursing home residents. Characterization and measurement. *Am J Geriatr Psychiatry* 2000; 8:75-83.
15. Iverson GL, Hopp GA, DeWolfe K, et al. Measuring change in psychiatric symptoms using the Neuropsychiatric Inventory: Nursing Home version. *Int J Geriatr Psychiatry* 2002; 17:438-443.
16. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Res* 1991; 38:143-162.

17. Lueken U, Seidl U, Völker L, et al. Development of a short version of the Apathy Evaluation Scale specifically adapted for demented nursing home residents. *Am J Geriatr Psychiatry* 2007; 15:376-385.
18. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011; 11: doi:10.1186/1471-244X-11-91.
19. Kat MG, de Jonghe JF, Aalten P, et al. [Neuropsychiatric symptoms of dementia: psychometric aspects of the Dutch Neuropsychiatric Inventory (NPI)]. *Tijdschr Gerontol Geriatr* 2002; 33:150-155. In Dutch.
20. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiatry* 1988; 23:271-284.
21. Dröes RM. The Dutch version of the Cornell scale for depression in dementia. In: Dröes R, editor. *Amsterdamse Ontmoetingscentra, een nieuwe vorm van ondersteuning voor dementerende mensen en hun verzorgers.* Nijkerk: Thesis Publishers; 1996, p.157.
22. Alexopoulos GS, Abrams RC, Young RC, et al. Use of the Cornell scale in nondemented patients. *J Am Geriatr Soc* 1988; 36:230-236.
23. Barca ML, Engedal K, Selbaek G. A reliability and validity study of the Cornell scale among elderly inpatients, using various clinical criteria. *Dement Geriatr Cogn Disord* 2010; 29:438-447.
24. Leontjevas R, van Hooren S, Mulders A. The Montgomery-Asberg Depression Rating Scale and the Cornell Scale for Depression in Dementia: A validation study with patients exhibiting early-onset dementia. *Am J Geriatr Psychiatry* 2009; 17:56-64.
25. Cronbach LJ, Warrington WG. Time-limit tests: Estimating their reliability and degree of speeding. *Psychometrika* 1951; 16:167-188.
26. Jackson S. *Research methods and statistics: A critical thinking approach*, Belmont CA: Wadsworth Cengage Learning, 2008.
27. Hanley JA,McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143:29-36.
28. Zhou X, Obuchowski NA,McClish DK. *Statistic methods in diagnostic medicine*, New York: Wiley, 2002.
29. Hanley JA,McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148:839-843.
30. Van der Wurff FB, Beekman AT, Comijs HC, et al. [Apathy syndrome: A clinical entity?]. *Tijdschr Gerontol Geriatr* 2003; 34:146-150. In Dutch.
31. Streiner DL, Norman GR. *Health measurement scales: A practical guide to their development and use*, 2nd Ed. Oxford: Oxford University Press, 1995.
32. Gil R, Arroyo-Anllo EM, Ingrand P, et al. Self-consciousness and Alzheimer's disease. *Acta Neurol Scand* 2001; 104:296-300.

Results section

Chapter

6

Process evaluation to explore internal and external validity of the 'Act in case of Depression' care program in nursing homes

Ruslan Leontjevas, Debby L. Gerritsen, Raymond T.C.M. Koopmans, Martin Smalbrugge, Myrra J.F.J. Vernooij-Dassen

Journal of American Medical Directors Association, 2012, 13(5): 488.e1-8.

Abstract

Background: A multidisciplinary, evidence-based care program to improve the management of depression in nursing home residents was implemented and tested using a stepped-wedge design in 23 nursing homes: Act in case of Depression (AiD).

Objective: Prior to effect analyses, to evaluate AiD process data on sampling quality (recruitment and randomization, reach) and intervention quality (relevance and feasibility, extent to which AiD was performed), which can be used for understanding internal and external validity. In this paper, a model is presented which divides process evaluation data into first and second-order process data.

Methods: Qualitative and quantitative data based on personal files of residents, interviews of nursing home professionals, and a research database were analysed according to the following process evaluation components: sampling quality and intervention quality.

Setting: Nursing home.

Results: The pattern of residents' informed consent rates differed for dementia special care units and somatic units during the study. The nursing home staff was satisfied with the AiD program and reported that the program was feasible and relevant. With the exception of the first screening step (nursing staff members using a short observer-based depression scale), AiD components were not performed fully by NH staff as prescribed in the AiD protocol.

Conclusion: Although NH staff found the program relevant and feasible and was satisfied with the program content, individual AiD components may have different feasibility. The results on sampling quality implied that statistical analyses of AiD effectiveness should account for the type of unit, whereas the findings on intervention quality implied that, next to the type of unit, analyses should account for the extent to which individual AiD program components were performed. In general, our first-order process data evaluation confirmed internal and external validity of the AiD trial, and this evaluation enabled further statistical fine tuning. The importance of evaluating the first-order process data before executing statistical effect analyses is thus underlined.

Introduction

Depression is a serious health problem affecting 6 to 50% of nursing home (NH) residents.[1,2] It is associated with increased mortality,[3-5] low quality of life,[6-8] increased hospitalization rate and the use of health care services in NHs.[9,10] Unfortunately, depression is often unrecognized and is consequently under treated in NH residents.[2,11-14] Moreover, although collaborative care and additional psychotherapy provide more benefits for patients with depression than pharmacotherapy alone,[15] the use of antidepressants alone is the most common depression treatment in the NH population.[16,17] Considering that recognition and treatment of depression constitutes a quality indicator in NHs,[18] the Nijmegen University Network of NHs (UKON, www.uko-n.nl) developed an evidence and practice-based care program 'Act in Case of Depression' (AiD, in Dutch 'Doen bij Depressie'). This care program includes a structural multi-step assessment procedure (screening and diagnosing), and multidisciplinary stepped-care treatment protocols based on current Dutch and international guidelines[18,19] and evidence from research. AiD aims to decrease the depression prevalence and to improve the quality of life of NH residents by improving depression assessment and treatment. To investigate the (cost) effectiveness of the AiD program, a two-year randomized controlled trial (RCT) was performed.[20]

RCTs are widely considered as the 'gold standard' for providing credible evidence of effectiveness of a health program. However, study design alone should not be the main criterion for evidence credibility.[21] If a study possesses no internal validity due to an insufficient sample size or poorly performed program activities, its results will be considered as not credible by decision makers. Without internal validity, effect analyses will be meaningless, and a waste of efforts and resources. Thus, the evaluation of sampling quality (the recruitment and maintenance of participants) and intervention quality (the extent to which the program was performed) is essential *prior to* executing outcome analyses. Appraisal of both sampling and intervention quality is also important for clinicians, policymakers and other research consumers interested in the generalization and applicability of findings, which form external validity indicators.[22,23] For instance, recruitment rates lower than 10% are often found in RCTs measuring effectiveness in clinical practice,[24] and these low rates can influence the generalizability of the observed effects. The acceptability of the program by health professionals is another intervention quality indicator that can influence implementation of the program in practice.[22,24] Thus,

next to a proper design, evaluation of both sampling quality and intervention quality are essential for credibility of the results.

Process evaluation can be used for the appraisal of sampling and intervention quality. To date, there is no 'gold standard' for conducting a process evaluation, and different comprehensive frameworks have been proposed.[25-28] Next to sampling and intervention quality issues, such frameworks cover numerous qualitative and quantitative elements that can be used for different aims. For example, different barriers and facilitators have been reported for intervention research in NHs,[29] and process evaluation can reveal such factors for an individual health intervention. Detailed documentation of strategies used by the implementers is important for replication and for future transition from research evidence into health practice.[27,30] These elements are not only important for the accumulation of implementation knowledge, but also for better understanding of why the intervention was or was not successful.[27] However, we believe that the elements that are important for the accumulation of implementation knowledge can be evaluated after performing the effect analyses and, therefore, are not of the same order as process data on sampling and intervention quality. The latter should be evaluated at an earlier stage to provide essential information about credibility and generalizability of health care interventions for decision makers.

Prioritizing the type and amount of process data can be important because time constraints and budgetary considerations may delay (or impede) proceeding according to one of the comprehensive frameworks.[25] Based on the framework of Linnan and Steckler.,[25] and the proposed criteria for assessing internal and external validity,[24,31] we present a model with first and second-order process data (Figure 6.1). First-order process data address sampling and intervention quality, and second-order data address implementation knowledge.

We apply this model using the AiD program in order to evaluate first-order process data about sampling quality and intervention quality *prior to* effect analyses. By focusing on sampling quality, data can be provided about the sample at baseline and follow-up measures, which is important for both internal (e.g. was randomization adequately performed?), and external validity (e.g. was the sample representative?). By focusing on intervention quality, data can be provided for understanding internal and external validity, and for further fine-tuning (or even annulling) of the effect analyses.

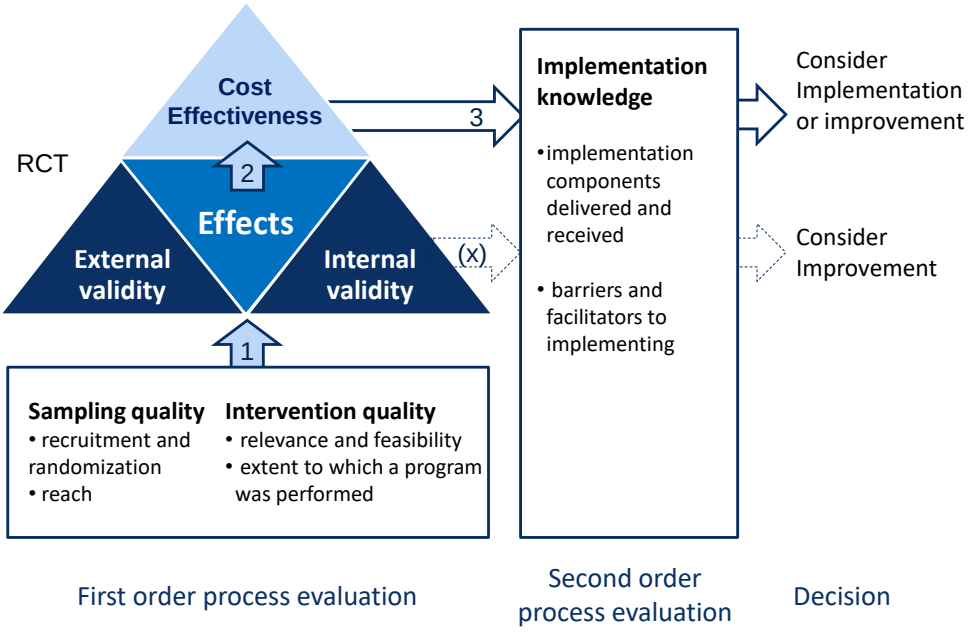


Figure 6.1 First and second order process data evaluation in relation to cost effectiveness analyses of a health program.

For confirming internal and external validity, first order process data should be analyzed (1). If first order process data confirms study validity, the results of (2) effect evaluation and cost-effectiveness analysis may be used, together with (3) implementation knowledge, to consider further implementing or improvement of the care program. If a study possesses no internal validity, or external validity is very limited, evaluation of effects could be terminated (X-path), or second order process data may be used for care program improvement.
Elements based on the framework of Linnan and Steckler:[25]
Recruitment and randomization: recruitment and randomization procedures, and barriers to the recruiting.
Reach: absolute or relative numbers of participants involved in the intervention.
Relevance and feasibility: indicators of program quality concerning feasibility, applicability and generalizability.
Extent to which a program was performed: measurement for the complex program elements performed per cluster and/or participant.
Implementation components delivered and received: the number of implementation components (e.g. study material, feedback) provided to and received by program performers (NH professionals).
Barriers and facilitators to implementing: factors at individual, organizational and environmental levels that (could) have influenced implementation of the program.

AiD is a complex program with various elements performed by a multidisciplinary team. If a certain program element was not performed adequately, the results could be attributed to program elements performed, rather than to the whole program (internal validity). Consequently, in further analyses, we should account rather for the extent to which the individual elements were performed than run an additional per-protocol analysis. If a

certain element was not performed or its acceptability by NH staff was low, the generalizability of the program can be questioned (external validity).

In this paper, two research questions regarding first-order process data of the AiD trial are addressed in order to assess internal and external validity of the intervention: what were (1) the sampling quality (2) and the intervention quality?

Methods

The process evaluation was conducted alongside the effectiveness multicenter study of the care program. Detailed information on the methods and the program are presented elsewhere.[20]

	Intervention group					Units [N]		
	1	2	3	4	5	Condition		Total
Measurement [N]	[7]	[6]	[7]	[6]	[7]	control	Interv.	
T0	c/s	c	c	c	c	[32*]	[0]	[32]
T1	i	c/s	c	c	c	[26]	[7]	[33]
T2	i	i	c/s	c	c	[20]	[13]	[33]
T3	i	i	i	c/s	c	[13]	[20]	[33]
T4	i	i	i	i	c/s	[7]	[26]	[33]
T5	i	i	i	i	i	[0]	[33]	[33]

N units summarized per condition

Figure 6.2 Graphical representation of the stepped wedge design with 6 measurements (T0-T5) and five intervention groups

[N]: number of nursing home units; c, control: measurement of usual care, control condition; c/s: measurement of usual care followed by the start of the implementation of the AiD program; i, Interv.: represent intervention-condition;
Measurements T0-T5 are approximately 4 months apart;
T0: baseline measurement, all groups are in the control condition;
T5: last follow-up, all groups are in the intervention condition;
32*: there were 32 units at T0, and 33 units at T1-T5, one unit that was randomized in group 5 was included at T1.

AiD study design

A stepped-wedge design with five clusters and six measurements (Figure 6.2) was used. This is a type of crossover design[32] in which NH units, randomly assigned to one of the clusters, cross over from the control to the intervention condition at different points in time, directly after a measurement (T0-T4, each 4 months apart). Consequently, at the last follow-up T5 the duration of the intervention condition varies from approximately 4

months to 20 months. In regular T0-T5 measurements, research staff unacquainted with the residents assessed scales on primary and secondary outcomes during an interview with the proxy (primary professional caregiver), and, if verbal communication was possible, with the resident.

Program components

Detailed information about the program is presented elsewhere.[20] In short, the AiD care program proposes five phases of depression management: (1) a screening procedure by the nursing staff using a short observer-rated scale;[33] (2) an extensive screening by a psychologist interviewing the client,[34] or, if the client cannot be interviewed, the professional caregiver[35]; (3) a diagnostic procedure by the psychologist and an elderly care physician; (4) treatment; and (5) monitoring of the treatment results using the instruments from phase 2. In phase 4, the multidisciplinary team (psychologist, physician, and care team including nursing staff and recreational therapist) can provide up to three treatment modules: (1) environmental and behavioral strategies, including a pleasant activities plan and a day structure program in case of depressive symptoms; (2) complementary to module 1, psychotherapy in case of depression; and (3) medication, especially if depression is severe. Each phase starts when indicated by the results of a preceding phase, with the exception of the first one, which needs to be completed each 4th month in the AiD intervention condition.

The research team used the following implementation strategy elements during the study: (1) provision of information and practical tools to NH professionals; (2) education consisting of: a 3 hour course about depression and the AiD components for nursing staff members, a 3.5 hour life review therapy training for the psychologists,[36] and provision of a medication protocol to the unit physician; and (3) tailored communication with psychologists and physicians about individual depression scores obtained during regular T0-T4 measures by the research team (the same used by the unit psychologists in phase 2) and, if available, depression assessment results (phases 1-3) and indicated treatment provided by the NH team.

First-order process data

To evaluate the sampling quality, *recruitment and randomization* was defined by description of (1) the recruitment and randomization procedure for nursing home units; (2) the informed consent and allocation procedure for the residents; and (3) barriers and facilitators to the recruiting of the residents. *Reach* was determined by the proportion of

the residents participating in the AiD study and the number of stakeholders involved in the intervention.

For the second research question, *intervention quality* was determined in two steps. To determine *relevance and feasibility*, nursing staff members were asked to rate on a 5-point scale (here and further from 1=totally disagree to 5= totally agree) whether the intervention raised their awareness of depression. The psychologists, physicians and unit managers were asked to what extent the AiD program was clear, and, both in general and specifically for their residents, relevant and feasible on a 5-point scale. They were also asked using open-ended questions to describe the advantages and disadvantages of the AiD program, and whether stakeholders would recommend the AiD program to other units or colleagues. Additionally, the overall satisfaction with program components, and program content and materials was assessed on a 10-point scale (0=completely dissatisfied, 10= completely satisfied).

Second, for the *extent to which the AiD program was performed,* the proportion of initiated cases of AiD components was determined per component in relation to indicated cases in residents of the unit.

Data collection procedure

Data were collected using a research database, residents' personal files, and structured phone interviews with physicians, psychologists and unit managers. Psychologists and physicians in intervention units were asked to fill in a digital form about clearness, relevance and feasibility of the program. Further, structured interviews of one hour were held at T5 with physicians, psychologists and unit managers. The responses to open-ended questions in interviews (by phone and on site) were typed by the interviewer (RL or a research assistant), and checked for accuracy during the interview. The first and second authors assigned conceptual labels to textual fragments. Disagreements were solved by discussion. Codes referring to the same phenomena/issues were grouped by RL and DG into categories.

Results

Sampling quality appraisal

Recruitment and randomization

Unit-level: The UKON steering committee invited the collaborative NH organizations to participate in the AiD study by providing one dementia special care unit (DSC) and one

somatic unit per NH where no other trials were held. Exclusion criteria were large-scale reorganization or rebuilding activities. In total, 17 somatic units and 16 DSC units from 23 NHs were proposed by 10 of the 11 organizations (one organization did not participate due to an ongoing reorganization). Units varied in size (Mean=28 residents, SD=9.7): eight units had between 14 and 20 residents; 12 units between 21 and 30 residents; 11 units between 31 and 34 residents; one DSC unit had 50 residents, and one somatic unit had 62 residents. All units participated until T5 (end of the study), and implemented at least one AiD phase. One DSC unit (group 5) discontinued program implementation due to relocation (T5 measurements were completed 2 months earlier than planned), although phases 1 to 3 were completed. To avoid contamination inside NHs, units were randomized per NH (in pairs if belonging to one NH). One DSC unit enrolled at T1 due to delayed inclusion, and it formed a pair with another unit from the same NH. For NH randomization, computer-generated random numbers were used by the researcher (RL) who was not involved in the recruitment of units or residents.

Resident level: Informed consent (IC) forms and instructions were provided by the research team to the unit manager. Nursing staff was responsible for approaching the residents to request IC. Before T0, the staff distributed IC forms and information brochures to all unit residents or their legal representatives (if resident was incapable of giving IC). To determine residents' capability of giving IC, nursing staff used residents' legal competency records. In case of doubts, legal representatives were to be approached. Legal representatives provided IC for 22% somatic residents, and for 99% DSC residents. Information about neither the precise schedule of the unit cross over to the intervention condition nor concrete program components was given to the residents or their representatives. Guided by ethical considerations to respect privacy, staff and residents were assured that no information would be collected about those residents who declined participation. Other than no provision of written informed consent, no additional exclusion criteria were used. The research team opened a phone help line for questions about the procedure and the program. After each measurement, they contacted unit managers by phone and by e-mail to remind them to attempt to recruit newly admitted residents (by exception, only e-mail was used after T4). Participation was voluntary and residents and/or their representatives did not receive any financial reward, nor was NH staff rewarded for performing the IC procedure.

Barriers and facilitators to recruiting: Ten of the 32 NH unit managers interviewed at T5 (one manager led two units) named prejudice about depression as being a reason to withhold the IC. Many residents or their representatives would not provide consent

because residents 'did not have depression' and 'intervention was not needed.' Other barriers included low motivation of some clients or their representatives, client's poor health, and relatives' concerns about resident burden, or worsening care due to staff time spent on the study. High staff workload and priorities other than the IC procedure were barriers for nursing staff. When asked about facilitators, the managers mentioned the tools and information provided by the research team (brochures, letter samples, forms, phone help line and personal contact on request), and personal contact with the residents/representatives by the nursing team.

Table 6.1
Results of recruitment of the residents for the Act in case of Depression study (AiD)

	Total	Per condition		Per type of unit	
		Control	Intervention	DSC	Somatic
Units, [N], T0-T5	[32-33]	[32-0]	[0-33]	[15-16]	[17-17]
Unit size, Mean n(SD)	28(10)	n/a	n/a	27(8)	28(11)
Residents:					
- n, T0-T5	886-902	886-0	0-902	413-436	473-466
- total IC, n	883	n/a	n/a	451	432
- deceased,n/%	331/37	176/ n/a	155/ n/a	197/44	134/31
- moved, n/%	108/12	32/ n/a	76/ n/a	38/9	70/15
- IC withdrawn, n/%	12/1	2/ n/a	10/ n/a	3/1	9/2
Per measurement [N] n/%(SD)					
T0* [32]	488/57(15)	[32] 488/57(15)	[0] n/a	254/62(15)	234/52(14)
T1 [33]	521/59(16)	[26] 410/58(15)	[7] 111/64(21)	279/65(18)	242/53(12)
T2 [33]	513/57(15)	[20] 310/57(15)	[13] 203/57(17)	263/61(17)	250/54(13)
T3 [33]	493/56(17)	[13] 193/55(16)	[20] 300/57(18)	250/58(20)	243/53(14)
T4 [33]	498/56(14)	[7] 98/54(7)	[26] 400/57(16)	245/57(16)	253/56(12)
T5 [33]	432/49(15)	[0] n/a	[33] 432/49(15)	213/50(16)	219/49(15)

Per condition: data are presented per control and intervention conditions. All clusters were in the control condition at T0 and crossed over to the intervention condition after one of the measurements (Figure 6.2) resulting at T5 in [33] units in the intervention condition (the last 7 units crossed over after T4).
Per type of unit: the same data as in 'per condition' but now presented for dementia special care (DSC) and somatic units.
deceased, moved and IC withdrawal, n/% : number and proportion of residents recruited.
Per measurement, % (SD): mean and standard deviation representing proportions of residents with an IC per unit; T0*: 32 units at T0 and 33 at T1-T5, one DSC unit was randomized in group 5 and included after T0.
Abbreviations, [N]: number of units; n: number of residents within the units; n/a: not applicable (all units cross over from control to intervention condition); IC: informed consent; total IC, n: number of original IC received.

Reach

Residents: Before T0, IC was provided by 547 (62%) residents or their representatives. 39 (7%) residents died, 17 (3%) moved to another unit and three (1%) withdrew IC before T0 (not shown). In total, 44% of 451 recruited DSC residents and 31% of 432 somatic residents died during the study (Table 6.1). Residents' participation rate varied per unit and per measurements T0-T5. In total, DSC units had higher rates than somatic units. With the exception of T0 and T5, the rates were virtually similar for the units during the control and intervention conditions. The last follow up at T5 showed a drop of 7% for both DSC and somatic units.

Stakeholders: In the intervention condition, 85% (N=712) of the employed nursing staff members attended the course. 581 of them (82%, range 42-100% per unit) were still employed at T5. 49 physicians, 42 psychologists, and 44 unit managers were involved in the intervention condition with 22, 19 and 11 cases of employee replacement, respectively. Although some stakeholders were involved simultaneously in two different units, and some (N=9) transited from one unit to another, contamination bias did not occur.

Intervention quality appraisal

Relevance and feasibility: According to nursing staff members, their awareness of depression in NH residents was raised due to the intervention (N=353, Mean=3.4, SD=0.8). The main stakeholders (N=95) evaluated the program as clear (N=88, M=4.0, SD=0.6), fairly well feasible in nursing homes (N=89, M=3.5, SD=0.7), and relevant for NH residents (N=91, M=4.0, SD=0.8). They tended to evaluate the AiD as less feasible (N=88, M=3.3, SD=0.9) for their own unit, and less relevant for their unit residents (N=92, M=3.8, SD=0.9) than in general.

When asked about the disadvantages of the AiD program, 39 (41%) stakeholders mentioned additional staff time needed to implement the program. Also mentioned were: feasibility problems (6%), risk of over-diagnosing of depression (3%), puzzling texts (3%), and standardization (3%). Regarding advantages of the AiD program, stepped-care standardization (46%) was mentioned most often, followed by the applicable program tools (29%), and the ability to improve staff knowledge and awareness of depression (28%). A time component, such as regularity and 'more prompt' recognition of depression was mentioned by 14 (15%) stakeholders. Other advantages of using AiD were: 'client-centred approach' (10%), a better collaboration between the disciplines (14%), theory-based phases (8%), and less medication use (4%). Most of the stakeholders, would

recommend the program as it is (53%) to other colleagues, or at least some elements of it (33%). 3 physicians would not recommend the program to a colleague because of time investment required (N=2), or irrelevance for a physician (N=1). The following examples of successful implementation were given: improved staff awareness (N=52), better collaboration (N=43); individual AiD phases performed and, especially, treatment module 1 containing a pleasant activities plan and a day structure program (N=34); program structure implemented and standardised procedures (N=24), and positive effects on clients (N=15). Regarding the period before the AiD, no use of a standardized depression program was reported in any of the units, and, according to the stakeholders, depression was mostly assessed after possible depressive complaints were raised by the nursing staff, the resident or next of kin.

In general, the stakeholders were satisfied with the program material presentation and its content (Table 6.2). The overall program satisfaction score was 7.5 (scale 0 to 10), and the score for satisfaction with the support provided by the research team was 8.1.

Table 6.2
Satisfaction of the main stakeholders with the program's material presentation, program content, and overall satisfaction with the program. Scale 0 (totally dissatisfied) to 10 (totally satisfied). Number of persons interviewed: 31 physicians, 33 psychologists and 32 unit managers.

	Presentation material	Program content	Overall Program
Physicians, M(SD)	7.9 (0.7), n=26	7.9 (0.4), n=25	7.5 (0.8), n=26
Psychologists, M(SD)	7.8 (0.8), n=33	7.8 (0.6), n=32	7.6 (0.6), n=26
Unit Managers, M(SD)	7.9 (0.8), n=29	7.9 (0.7), n=27	7.4 (0.8), n=28
Total, M(SD)	7.9 (0.8), n=88	7.8 (0.6), n=84	7.5 (0.7), n=80

Missing values due to 'not applicable' or 'don't know' answers.
At T5, 32 physicians, 31 psychologists and 32 unit managers represented 33 units. Of them, 3 physicians refused the final interview due to other priorities. Additionally, 2 physicians and 2 psychologists who switched jobs or units between T4 and T5 were also interviewed.

Extent of program performance: Of 883 residents involved in the study, phase 1 could not be completed for 243 (28%) residents due to: death (N=161), relocation (N=57), recruitment after the last screening cycle was completed (N=22), and participation withdrawal (N=3), which was between 3 (group 1) and maximum 23 months (group 5) prior to program implementation within the unit.

Table 6.3
Extent to which AiD phases were performed as indicated

	Indicated	Initiated	All	Per intervention group					Per type of unit	
	[N]	[N]/%	units	1	2	3	4	5	DSC	Som
Units, [N: DSC, Somatic]			[16,17]	[4,3]	[3,3]	[2,5]	[4,2]	[3,4]	[16]	[17]
Crossing over after				T0	T1	T2	T3	T4		
AiD phase										
Depression assessment										
- Ph. 1 Detection	[33]	[33]/100	86(16)	82(3)	76(14)	95(9)	85(23)	89(20)	82(17)	89(14)
- Ph. 2 Screening	[33]	[27]/82	55(36)	29(24)	58(25)	72(24)	56(46)	63(47)	47(39)	64(32)
- Ph. 3 Diagnosing	[31]	[23]/74	52(41)	27(34)	59(38)	54(40)	49(48)	77(41)	52(39)	53(43)
Total assessment	[33]	[33]/100	76(18)	69(9)	70(12)	87(12)	74(26)	78(24)	69(19)	82(15)
Treatment										
- Ph. 4 module 1	[32]	[20]/ 63	42(40)	39(36)	48(40)	46(41)	43(48)	32(47)	48(42)	35(38)
- Ph. 4 module 2	[29]	[12]/41	33(45)	16(37)	56(50)	40(55)	25(39)	35(49)	17(35)	49(48)
- Ph. 4 module 3	[30]	[15]/50	40(46)	45(46)	17(41)	58(49)	53(52)	25(43)	44(45)	37(48)
Any treatment module	[32]	[23]/72	40(36)	30(22)	44(36)	54(42)	43(38)	30(44)	43(33)	38(40)
Monitoring (Ph. 5)	[31]	[11]/35	18(30)	15(20)	25(39)	29(41)	17(26)	0(0)	22(32)	14(27)

Indicated [N]: number of units with residents for whom an AiD component was indicated;
Initiated [N] / %: number of units that initiated an AiD component for their residents, and percentage of [N] indicated;
Mean proportion % (SD): proportion of initiated cases of AiD components determined in relation to indicated cases in residents of the unit, reported as mean percentage and standard deviation (SD) for the units per intervention groups, and for DSC and somatic units.
Total depression assessment and total treatment scores were calculated based on the number of components the resident received in relation to the number of components that were indicated.
Abbreviations, [N]: number of units; DSC: dementia special care; Som: somatic; Ph.: phase.

Phase 1 was performed in all units, in most of the indicated cases (mean proportion of initiated cases when indicated was 86%, Table 6.3). Phase 5 was performed the least often (18%), and, in total, only 11 out of 31 units that should have performed this phase actually did perform it. Table 6.3 shows that the extent to which the units performed the components varied in the intervention groups with the largest variation found for treatment modules (*SD* up to 55%). In total, depression assessment phases 1, 2 and 3 were performed almost twice as often (76%) as treatment modules (40%). Psychotherapy (module 2) was performed in somatic units (49%) almost three times as often as in DSC units (17%), whilst there was an indication that behavioral strategies (module 1) were performed more often in DSC units (48%) than in somatic units (35%).

Discussion

In this paper, we illustrated a model that can be used for ordering process data. In addition to receiving input for our own effect analyses, the results described in this paper are important for improving NH depression management and NH research in general.

Sampling quality

The process evaluation showed a different pattern for the provided informed consent (IC) rate in 16 DSC and 17 somatic units. Prior to the baseline measurement (T0), DSC units had a higher IC rate than somatic units. However, nursing staff members of somatic units succeeded in gradually improving the rate by approaching newly admitted residents or their representatives before the fifth measurement (T4), 20 months later; whereas DSC units showed a gradual decline resulting in a comparable rate at T4. Process data implied that the IC rate would drop dramatically without an ongoing recruitment of the newly admitted residents: nearly one out of two recruited DSC residents, and one out of three somatic residents died, and a significant portion of the residents (DSC: 9%, somatic: 15%) moved to another unit during the two-year study.

The rates of residents not providing consent were higher than the expected negligible in DSC, and 20% in somatic units, which were based on Dutch observational studies in NHs.[20] Comparison of these rates with rates in other NH depression studies is difficult. In many such studies, reach and recruitment have seldom been fully described. In a study of Brodaty et al.,[37] researchers obtained consent for a comparable proportion of 56% of the residents with dementia before they further applied other exclusion criteria. For residents without cognitive impairments, Rosen et al.[38] reported a comparable proportion (50%) of residents who agreed to diagnostic evaluation in long-term care facilities. For depression and anxiety screening in Dutch residential homes for the elderly, Dozeman et al.[39] reported 44% consent.

Although external validity could be influenced by the units' convenience sampling and decline in the IC rate at the end of the study, bias in recruitment of residents was unlikely because recruiters did not have a substantial impact on who was recruited.[24] According to the unit managers, poor state of health was one of the reasons residents or their representatives withheld IC, and this could influence external validity too. On the other hand, we did not apply additional exclusion criteria for residents, and that should improve external validity. With regard to internal validity, it was unlikely to have been influenced by the sampling quality, as units did not differ in the recruitment strategy used, and all of them crossed over according to the stepped-wedge design from the control to the

intervention condition. However, further analyses should account for the unit type, as DSC and somatic units differed in the IC rates.

In addition, future depression studies in NHs should account for barriers to recruiting, such as reluctance to participate in a study on depression, and residents' or their representatives' concerns about extra burdening, or less appropriate care during the study due to expected prioritizing trial activities over daily care. Personal communication about the study, and well organized tools such as brochures and a phone help line are important in this regard: they were appreciated by NH teams, participants and their representatives.

Intervention quality

Relevance and feasibility: With regard to relevance and feasibility, the study showed stakeholders' satisfaction with the program. Most of the stakeholders would recommend the program as it is, or some of its individual elements to other colleagues, which underlines the relevance of the program. Nursing staff members and main stakeholders (physicians, psychologists, and unit managers) alike reported that depression awareness, and knowledge about depression improved in nursing staff. This is encouraging, as recognition of depression by nursing staff is low in NHs,[40] and those who have the most daily contact with residents have limited knowledge about depression.[11]

The relevance and feasibility reported by the stakeholders implies that the program has good external validity. However, some disadvantages of the program were also mentioned. The need for additional time to implement the program was a disadvantage according to 41% of stakeholders, which is not surprising as 'insufficient time on the job to implement new ideas' was one of the most frequently named items in 63 studies into nursing interventions reviewed by Kajermo et al.[41] It is possible that stakeholders experienced a shortage of time for program implementation, and, therefore, evaluated feasibility and relevance of the program in their own units lower than that in the NH setting in general. Future studies should consider that the intervention quality evaluation by stakeholders may differ when reported from different perspectives: from the perspective of their own organization or unit, and in the NH setting in general.

Extent of program performance: Considering program performance, the results showed that adherence to the program varied for the different AiD phases. Remarkably, nursing staff members of all 33 units carried out phase 1 at least once for most of their residents. The successful application of phase 1 is an important finding for both internal and external validity, and can be important for depression management and research in NHs: a short universal screening scale for both somatic and DSC residents was chosen to

facilitate the screening,[33] and its regular use was feasible in the study. Given this success, and lower figures for phase 2, additional analyses could be performed on whether using phase 1 and performing a comprehensive diagnostic procedure without extensive interview by the psychologist in phase 2 (screening) may improve AiD feasibility and still be effective.

It is worrying that treatment and monitoring according to the AiD program were performed less often than the previous three assessment phases. Treatment module 2 was less frequently used in DSC units than in somatic units, which may reflect existing challenges associated with psychotherapy in dementia. However, AiD module 2 also provides psychologists with a protocol to intervene through the nursing staff (mediative behavior therapy) when a 'talk therapy' with the resident is not possible.[20] Therefore, it is important to explore whether AiD mediative components need to be improved. On the contrary, module 1 (pleasant activities plan) was performed less often in somatic units than in DSC units, which may reflect characteristics of somatic residents (e.g. more physical disabilities that hinder the participation in activities), or organizational issues (e.g. less activities suitable for somatic residents). In general, more research on how protocols and procedures on depression treatment and monitoring can be improved is needed in order to facilitate depression management in NHs, and AiD in particular.

With regard to intervention quality we conclude that, although the stakeholders reported the program was relevant and feasible, feasibility, which is an important criterion for external validity, may not be the same for all phases. Considering that some components were not performed in all indicated cases, and there was an indication that intervention groups differed in performing the components, statistical analyses should account for this finding affecting internal validity. In addition to intention-to-treat analyses, the possible effects of individual phases should be explored. At least, when exploring the effects of treatment, the analyses should also account for assessment phases because they were performed the most often. Without performing the AiD treatment components, depression assessment might induce changes in treating the residents by NH staff (non-specific effect). Differences in performing the components between DSC and somatic units underpin, again, the importance of accounting for the type of unit in analyses.

Strengths and weaknesses

To the best of our knowledge, the AiD trial on the effects of a multidisciplinary care program is the largest RCT on depression recognition and treatment in NHs to date. The type of design used, a stepped-wedge design, contributes to the uniqueness of the study. For the process evaluation, different sources were used, and virtually all physicians,

psychologists and unit managers available at the end of the study were interviewed. However, the main stakeholders often changed their jobs or units, and it is known from previous research that staff turnover could be a barrier to implementation of health programs,[42] and research in NHs.[43] Therefore, exploration into whether staff changes during the study affected the implementation may be needed to determine why individual components were not implemented in all units.

Several limitations of our study are worth noting. The reasons why a certain unit was provided by the NH were not assessed. Furthermore, due to ethical considerations, we did not collect information about residents in the units that were not provided, and we did not collect data on residents who gave no consent within provided units. Therefore, we cannot compare our sample with non-participants. Future analyses on the sample characteristics in the effect study can reveal whether our sample was representative for the Dutch NH population and comparable with samples in other studies regarding demographic characteristics and outcomes. Furthermore, in order not to influence the residents, or compromise their anonymity and/or the anonymity of their representatives, the IC procedure was performed by the nursing staff, and not by the research team. Consequently, the staff and not the residents and their representatives provided us with information about the reasons in cases where IC was withheld. A recall bias was possible. And last, although the first-order process data analysis is important for confirming internal and external validity of the study, it does not mean that no further validity appraisal is needed. This first-order process data evaluation gives the go-signal for further analyses that should also include (various types of) validity appraisals. Second-order process data evaluation is needed to determine whether and why certain implementation strategy elements (e.g. education) influenced the program performance and the intervention results.

Conclusions

The results on the sampling quality implied that effect analyses should account for the type of unit, whereas the findings on the intervention quality implied that, next to the type of unit, the extent to which individual AiD program components were performed should be taken into account. Although NH staff found the program relevant and feasible and was satisfied with the program content, individual phases could have different feasibility. The first phase, including the use of a short observer rated depression scale by the nursing staff, was performed the most often and monitoring of the treatment results the least. In general, evaluation of the first-order process data confirmed the validity of the AiD trial,

and this evaluation enabled further statistical fine tuning. The importance of evaluating the first-order process data before executing statistical effect analyses is thus underlined.

Literature

1. Jongenelis K, Pot AM, Eisses AM, et al. In Dutch [Depression among older nursing home patients. A review]. *Tijdsch Geront Geriat* 2003;34:52-59.
2. Kramer D, Allgaier AK, Fejtkova S, et al. Depression in nursing homes: prevalence, recognition, and treatment. *Int J Psychiatry Med* 2009;39:345-358.
3. Barca ML, Selbaek G, Laks J, et al. Factors associated with depression in Norwegian nursing homes. *Int J Geriatr Psychiatry* 2009;24:417-425.
4. Katz IR, Lesher E, Kleban M, et al. Clinical features of depression in the nursing home. *Int Psychogeriatr* 1989;1:5-15.
5. Rovner BW, German PS, Brant LJ, et al. Depression and mortality in nursing homes. *JAMA* 1991;265:993-996.
6. Hoe J, Hancock G, Livingston G, et al. Quality of life of people with dementia in residential care homes. *Br J Psychiatry* 2006;188:460-464.
7. Wetzels RB, Zuidema SU, de Jonghe JF, et al. Determinants of quality of life in nursing home residents with dementia. *Dement Geriatr Cogn Disord* 2010;29:189-197.
8. Gonzalez-Salvador T, Lyketsos CG, Baker A, et al. Quality of life in dementia patients in long-term care. *Int J Geriatr Psychiatry* 2000;15:181-189.
9. Smalbrugge M, Pot AM, Jongenelis L, et al. The impact of depression and anxiety on well being, disability and use of health care services in nursing home patients. *Int J Geriatr Psychiatry* 2006;21:325-332.
10. Bartels SJ, Horn SD, Smout RJ, et al. Agitation and depression in frail nursing home elderly patients with dementia: treatment characteristics and service use. *Am J Geriatr Psychiatry* 2003;11:231-238.
11. Ayalon L, Arean P, Bornfeld H. Correlates of knowledge and beliefs about depression among long-term care staff. *Int J Geriatr Psychiatry* 2008;23:356-363.
12. Gruber-Baldini AL, Zimmerman S, Boustani M, et al. Characteristics associated with depression in long-term care residents with dementia. *Gerontologist* 2005;45:50-55.
13. Snowdon J, Fleming R. Recognising depression in residential facilities: An Australian challenge. *Int J Geriatr Psychiatry* 2008;23:295-300.
14. Damian J, Pastor-Barriuso R, Valderrama-Gama E. Descriptive epidemiology of undetected depression in institutionalized older people. *J Am Med Dir Assoc* 2010;11:312-319.
15. Oestergaard S, Moldrup C. Improving outcomes for patients with depression by enhancing antidepressant therapy with non-pharmacological interventions: a systematic review of reviews. *Public health* 2011;125:357-367.
16. Levin CA, Wei W, Akincigil A, et al. Prevalence and treatment of diagnosed depression among elderly nursing home residents in Ohio. *J Am Med Dir Assoc* 2007;8:585-594.
17. Dorland L, Pot A, Verbeek M, et al. In Dutch: [Psychische hulpverlening voor ouderen in verzorgings- en verpleeghuizen; deelstudie 7.] In Monitor geestelijke gezondheidszorg ouderen, rapportage 2006 Utrecht: Trimbos-Instituut, 2007.

18. AGS/AAGP. The American Geriatrics Society and American Association for Geriatric Psychiatry recommendations for policies in support of quality mental health care in U.S. nursing homes. *J Am Geriatr Soc* 2003;51:1299-1304.
19. Trimbos-instituut. In Dutch: [Addendum ouderen bij de multidisciplinaire richtlijn depressie]. Utrecht: Trimbos-instituut, 2008.
20. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011;11:91.
21. Rychetnik L, Frommer M, Hawe P, et al. Criteria for evaluating evidence on public health interventions. *J Epidemiol Commun H* 2002;56:119-127.
22. Glasgow RE, Klesges LM, Dzewaltowski DA, et al. Evaluating the impact of health promotion programs: using the RE-AIM framework to form summary measures for decision making involving complex issues. *Health Educ Res* 2006;21:688-694.
23. Katrak P, Bialocerkowski AE, Massy-Westropp N, et al. A systematic review of the content of critical appraisal tools. *BMC Med Res Methodol* 2004;4:22.
24. Eldridge S, Ashby D, Bennett C, et al. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ* 2008;336:876-880.
25. Linnan L, Steckler A. Process Evaluation for Public Health Interventions and Research: An Overview. In: Steckler A,Linnan L, eds. *Process Evaluation for Public Health Interventions and Research*. San Francisco, CA, US: Jossey-Bass, 2002.
26. Reelick MF, Faes MC, Esselink RA, et al. How to perform a preplanned process evaluation for complex interventions in geriatric medicine: exemplified with the process evaluation of a complex falls-prevention program for community-dwelling frail older fallers. *J Am Med Dir Assoc* 2011;12:331-336.
27. Hulscher ME, Laurant MG, Grol RP. Process evaluation on quality improvement interventions. *Qual Saf Health Care* 2003;12:40-46.
28. Baranowski T, Stables G. Process evaluations of the 5-a-day projects. *Health Educ Behav* 2000;27:157-166.
29. Mentes JC, Tripp-Reimer T. Barriers and facilitators in nursing home intervention research. *West J Nurs Res* 2002;24:918-936.
30. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362:1225–1230.
31. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365:82-93.
32. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;28:182-191.
33. Leontjevas R, Gerritsen DL, Vernooij-Dassen MJFJ, et al. The Nijmegen Observer-Rated Depression Scale for Detection of Depression in Nursing Home Residents. *Int J Geriatr Psychiatr* In press.
34. Yesavage JA, Brinka TL, Roseb TL, et al. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* 1983;17:37-49.
35. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23:271-284.
36. Bohlmeijer E, Steunenberg B, Leontjevas R, et al. In Ducth: [Dierbare Herinneringen: Protocol voor Individuele life-review therapie gebaseerd op autobiografische oefening]. Enschede: Universiteit Twente, 2010.
37. Brodaty H, Draper BM, Millar J, et al. Randomized controlled trial of different models of care for nursing home residents with dementia complicated by depression or psychosis. *J Clin Psychiat* 2003;64:63-72.

38. Rosen J, Rogers JC, Marin RS, et al. Control-relevant intervention in the treatment of minor and major depression in a long-term care facility. *Am J Geriat Psychiat* 1997;5:247-257.
39. Dozeman E, van Schaik DJ, van Marwijk HW, et al. In Ducth: [Screening for depression and anxiety in residential homes for the elderly]. *Tijdsch Geront Geriat* 2008;39:100-106.
40. Bruhl KG, Luijendijk HJ, Muller MT. Nurses' and nursing assistants' recognition of depression in elderly who depend on long-term care. *J Am Med Dir Assoc* 2007;8:441-445.
41. Kajermo KN, Bostrom AM, Thompson DS, et al. The BARRIERS scale -- the barriers to research utilization scale: A systematic review. *Implement Sci* 2010;5:32.
42. Francke AL, Smit MC, de Veer AJ, et al. Factors influencing the implementation of clinical guidelines for health care professionals: a systematic meta-review. *BMC Med Inform Decis Mak* 2008;8:38.
43. Mentes JC, Tripp-Reimer T. Barriers and facilitators in nursing home intervention research. *Western J Nurs Res* 2002;24:918-936.

Chapter

Effects of the nursing home multidisciplinary care program 'Act in case of Depression' including structural depression assessment and treatment pathways: A stepped-wedge cluster randomized controlled trial

Ruslan Leontjevas, Debby L. Gerritsen, Martin Smalbrugge, Steven Teerenstra, Myrra J.F.J. Vernooij-Dassen, Raymond T.C.M. Koopmans

submitted

Abstract

Context: Depression in nursing home residents is underrecognized and undertreated.

Objective: To determine effects of a systematic approach to nursing home (NH) depression management including depression assessment and treatment pathways.

Design, Setting and Participants: Stepped-wedge cluster-randomized trial between May 2009 and April 2011. 16 dementia special care (DSC) and 17 somatic NH units in four Dutch provinces were randomized in 5 intervention groups. 403 DSC and 390 somatic residents were recruited with no exclusion criteria except informed consent refusal.

Intervention: The new program 'Act in case of Depression' (AiD) prescribes pathways for structural assessment (a two-step screening procedure and a diagnostic procedure), multidisciplinary treatment (basic behavioral, psychotherapeutic, pharmacological), and monitoring of treatment effects. In the control condition, units provided usual care.

Main outcome, cluster level: Depression prevalence based on proxy-reports (Cornell Scale for Depression in Dementia, CSDD-score>7).

Secondary outcomes, cluster level: severe depression (CSDD-score>11), self-reported depression prevalence (8-item Geriatric Depression Scale, GDS8-score>2), and severe depression prevalence (GDS8-score>4); resident level: proxy-reported and self-reported quality of life (Euroqol-5D), and depressive symptoms (CSDD-score, GDS8-score).

Results: In somatic units, AiD reduced CSDD-depression prevalence (effect estimate -7.3%; 95% CI, -13.7% to -0.9%; Cohen's d, -0.6), and CSDD-score (-0.8; 95% CI, -1.4 to -0.1). In both somatic and DSC units, self-reported quality of life improved in the intervention condition (3.4; 95% CI, 0.5 to 6.3). Larger effect-sizes were associated with the use of assessment algorithms (endpoint Cohen's d, -1.1) but not treatment algorithms, which were used poorly. Intervention effect was found neither on self-reported depression nor on severe depression outcomes. Proxy-reported QoL worsened in the intervention condition (-2.0, 95% CI, -3.5 to -0.5), but improved if AiD treatment protocols were used (DSC: 4.1; 95% CI, -0.1 to 8.3; somatic: 8.2; 95% CI, 3.9 to 12.4).

Conclusions: The use of a multidisciplinary depression program including structural assessment procedures reduces depression prevalence in somatic units. AiD improves quality of life as reported by residents of DSC and somatic units. Further research is needed on improving the implementation of treatment protocols in the NH setting and on treatment of severe depression.

Introduction

Depression is the main factor contributing to the growing population of residents with a mental illness in nursing homes (NHs);[1] it is associated with increased mortaity,[2,3] increased hospitalization and use of other care services.[4,5] Unfortunately, structural depression assessment using validated instruments has seldom been performed in NHs,[6] and the condition often goes unrecognized and undertreated.[7,8] Although collaborative care with psychotherapy provides more benefits for patients with depression than pharmacotherapy alone,[9] most NH residents with depression are only treated with antidepressants.[6,10] Research shows that psychosocial interventions are effective for depression in NH residents,[11] whereas a collaborative approach combining medical, psychiatric and nursing interventions is effective in residents needing both psychiatric and nursing care.[12] Given the underrecognition of depression in NHs, adequate collaborative depression management comprising evidence-based treatment procedures should certainly include structural depression screening and diagnosing procedures (depression assessment).[13-15] However, it remains to be tested whether care programs combining both structural assessment and collaborative treatment protocols are effective in reducing depression prevalence in NH residents.

In Dutch nursing homes, the presence of multidisciplinary care teams which include an elderly care physician,[16] a psychologist, a recreational therapist and nursing staff offers ideal preconditions for collaborative depression care. The Nijmegen University Network of NHs (UKON, www.uko-n.nl), an alliance between 12 care organizations and the Department of Primary and Community Care of the Radboud University Nijmegen, Medical Centre, developed a multidisciplinary depression program 'Act in case of Depression' (AiD). AiD introduces pathways for standardized depression assessment and treatment protocols based on current national and international guidelines[15,17] and evidence from research on psychotherapy and pharmacotherapy.

We conducted a pragmatic trial to test the hypothesis that depression prevalence reduces in NH units after the introduction of a structural approach to depression management. Nursing care for residents with dementia in the Netherlands is provided in separate dementia special care (DSC) units, and AiD accounts for the differences between DSC and somatic residents. For example, assessment and psychosocial treatment procedures account for possible language problems in dementia. We aimed to determine whether the effect of a structural approach to depression management is different in the two types of units.

In contrast to an explanatory trial that tests an intervention under ideal or selected conditions and which is concerned with how and why an intervention works, a pragmatic trial is concerned with the question if the intervention is effective under real-life conditions.[18] Because of the pragmatic attitude, we evaluated process data prior to the effectiveness analysis and determined that internal and external validity were acceptable. NH professionals described the AiD program, which they used in the intervention condition, as feasible and relevant. However, we found that they used AiD assessment procedures to a larger extent than treatment pathways.[19] Because performing assessment procedures might induce changes in treating the residents by NH staff even when protocolized treatment was not performed, we aimed to determine whether AiD assessment procedures and treatment pathways were related to program effects.

Methods

Design

During May 2009 to March 2011, we conducted a multicenter, stepped-wedge cluster-randomized trial with six measurements (T0-T5),[20] which is a type of cross-over design with repeated measurements. A stepped-wedge design maximizes statistical power, and requires fewer clusters compared to a parallel cluster RCT.[21] To avoid contamination, and because the intervention was conducted on the NH unit level, the NH unit was the unit of randomization, intervention and primary analysis. The units were assigned to one of five intervention groups, which crossed over from the control to the intervention condition (introduction of the intervention program) at different points in time directly after a measurement (T0-T4, each 4 months apart, Figure 6.2). The researcher, not involved in the recruitment, randomized clusters using computer generated random numbers. If there were two units in one NH location, both were allocated to the same intervention group to avoid contamination bias. The residents were blinded with regard to the unit cross-over to the intervention condition. Research staff who administered the outcome instruments was blinded to treatment provided for the resident.

Setting and Participants

Detailed recruitment procedures and results are reported elsewhere (chapter 6).[19] A research coordinator (DG) recruited units between February and May 2009: 11 UKON organizations (about 5,000 DSC and 3,300 somatic beds) were invited to provide units, maximally one DSC and one somatic unit per NH location. Exclusion criteria for organizations were large-scale reorganization or rebuilding activities (one organization

was excluded). Before T0, residents of the included units were recruited by unit staff, which approached all unit residents or their legal representatives. To improve generalizability and to avoid selection bias, no exclusion criteria were used for residents with the exception of not providing informed consent (IC). Newly admitted residents were recruited until the final follow-up.[19] Each participating resident and/or legal representative (in case the resident was incapable of giving IC) received written and verbal information about the study.

Intervention

AiD describes algorithms for three components for depression assessment, three treatment modules, and monitoring of the treatment results (Appendix A1, Figure A1.1). The first screening step is completed each 4th month in the intervention condition. To reduce false negatives, NH professionals may use information other than results of standardized screening in order to initialize a diagnostic procedure. The psychologists and physicians may diverge from the content of individual treatment protocols if deemed necessary, but are requested to use the pathways for collaborative treatment.

The research team provided units in the intervention condition with program texts and practical tools, a 3.5 hour educational course about depression and AiD to the nursing staff, and a 3.5 hour training session to the psychologists about life review therapy.[22] A physician involved in the program development contacted the unit physician by phone to discuss a medication protocol.

No specific information about AiD was provided to NH staff and residents during the control condition. In the control condition, the units did not use a structural approach to depression management: depression was assessed after indications of possible depression were raised by nursing staff, residents, or next of kin,[19] and teams did not use multidisciplinary pathways for depression treatment.

Outcomes and assessments

Depression: The endpoint was depression prevalence assessed on the unit level based on the proxy-reported Cornell Scale for Depression in Dementia (CSDD).[23] The CSDD consists of 19 items, each rated as 0=absent, 1=mild or intermittent, and 2=severe. Depression prevalence was calculated per unit as the proportion of residents with a score >7.[23] A score >11 indicated severe depression[23] (secondary outcome).

Because proxy-reported reports of psychopathology including depression may differ from self-reports,[24-27] next to the *proxy-reported* CSDD outcomes, the 8-item Geriatric Depression Scale (GDS8)[28] was used for the assessment of *self-reported* depression

outcomes. The GDS8 was created by deleting items not applicable to NH residents from the original GDS.[29] A GDS8 score>2 indicated depression[28] and >4 severe depression.[30] On the resident level, the CSDD score and the GDS8 score were used as secondary outcomes for the severity of proxy-reported and self-reported depressive symptoms. **Quality of life:** Overall health-related quality of life (QoL) was assessed using a visual analogue 'thermometer' scale of the Euroqol-5 Dimensions.[31] A score of 0 represents the worst health state, while 100 represents the best. Because not all residents could be interviewed due to cognitive problems, the proxy-based approach was used for QoL. Because the self-report may be considered the gold standard for the accuracy of QoL measures,[27] residents who could respond were also interviewed for self-reported QoL using the Euroqol-5D.[31]

In sum, we used a proxy-reported and self-reported construct for four outcomes: prevalence of depression, prevalence of severe depression, QoL and depressive symptoms. Proxy-reported depression prevalence was the endpoint.

Proxies (primary professional caregiver) and residents were interviewed by the researcher, research assistant, and 32 graduate psychologists. Interviewers were trained in administering the scales and were not involved in providing AiD. The Mini-Mental-State Examination (MMSE)[32] was administered for global cognitive functioning (range 0 to 30, a higher score indicates less cognitive impairment). The extent to which the multidisciplinary teams used the algorithms (adherence rates) was conceptualized per unit as the proportion of the residents for whom NH staff used the algorithms for assessment procedures and the algorithms for treatment in relation to residents for whom the algorithms should be used. Assessment and treatment adherence rates were determined using residents' personal files, and structured phone interviews with physicians, psychologists and unit managers.

Study protocol change

The initial protocol prescribed the use of the CSDD in DSC units only and the GDS8 in somatic units.[20] The protocol was changed in the first week of T0 because, first, many residents in somatic units could not be interviewed due to language or cognitive problems resulting in missing GDS8 scores, and, second, the use of the same scale and the same scoring method was considered important for the comparison of the two unit types. The proxy-reported CSDD was the endpoint because, in contrast to the GDS8, the scale can be used in residents with and without cognitive impairments and the scale has been validated in dementia[23] and non-dementia[33] (as well as in NH residents[34,35]) .

Statistical analyses

Using the method described by Hussey and Hughes[36] and based on a significance level alpha of 0.05, a power of 0.8 and an ICC of 0.1, we calculated that 16 clusters were needed in a stepped-wedge trial for each unit type to allow multilevel analyses on depression prevalence.[20]

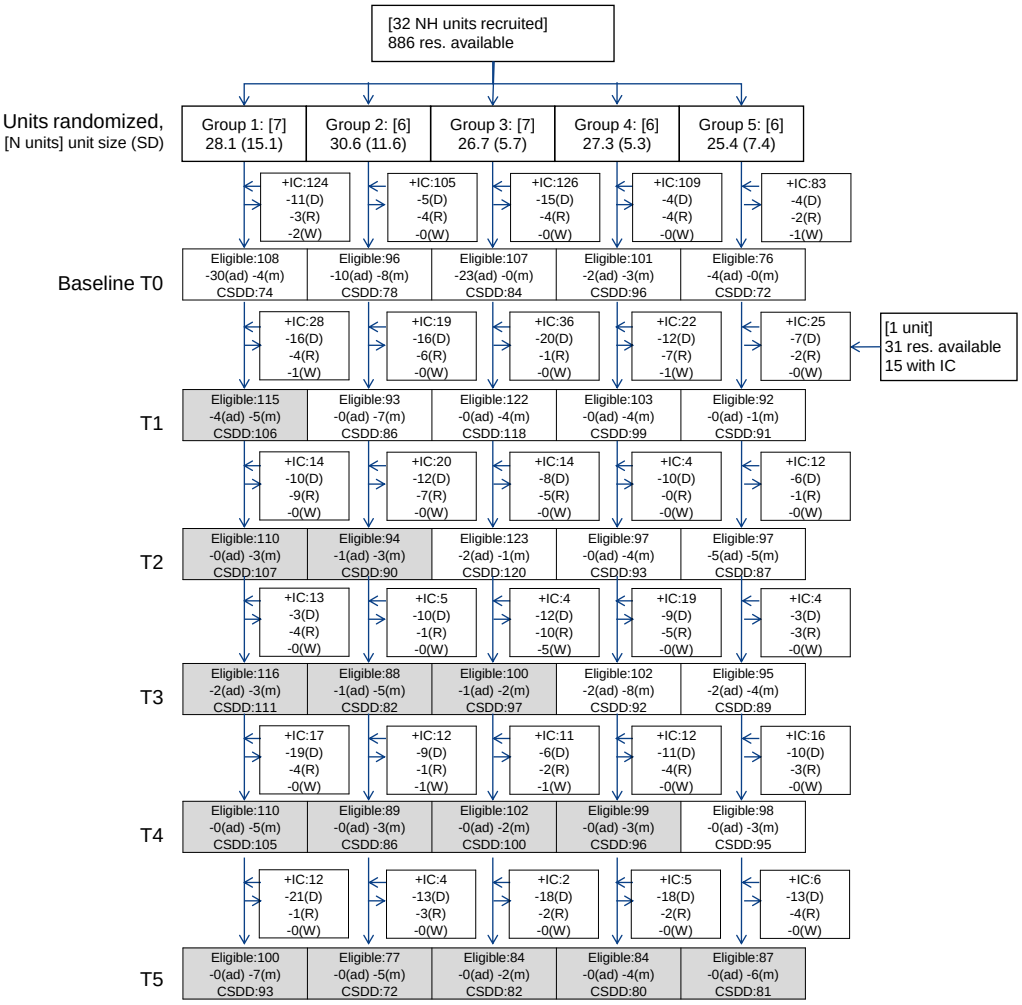
For descriptive analyses of unit and resident characteristics, we used SPSS 17.0.0 (Chicago, IL). To compare baseline characteristics and adherence rates in intervention groups, we used ANOVA. A t-test was used to compare DSC and somatic units. A Chi-square test was used for categorical variables of units. For a maximum of 4 missing individual CSDD items, a 0 score was imputed.[35] The same conservative imputation was performed for a maximum of two missing GDS8 items. Only residents scoring 15 or higher on the MMSE were included in GDS8 analyses.[28] Missing QoL observations were not imputed. To explore associations between self-reported and proxy-reported outcomes, partial correlations adjusted for age and gender were determined at entry of the resident in the study.

Since generalized linear mixed models did not converge, dichotomous variables (depression prevalence) were fit using linear mixed models with random effects for units and for subjects nested within units (SAS software 9.2, SAS Institute Inc., North Carolina). This is acceptable for dichotomous variables when degrees of freedom are sufficient.[37] The same method was used for continuous variables (QoL and depressive symptoms, resident level). All estimates were adjusted for age, gender, time trends (T0-T5), region of the country (province) and type of unit (somatic or DSC). To determine the intervention effect and to compare DSC and somatic units, likelihood ratio tests were used comparing a model with the main intervention effect and its interaction with type of unit, first, to a model without the interaction and, second, to a model without the main effect and interaction.[38] For the endpoint, Cohen's d is determined as the intervention effect divided by a standard deviation. Exploratively, we also tested the intervention effects on the GDS8 outcomes for all residents including those with low MMSE scores. Furthermore, the influence of the study duration and the intervention duration (number of 4-month periods between T measurements the resident was in the study, and in the intervention condition) was tested.

To determine whether adherence rates for depression assessment and for treatment protocols were related to the intervention effect, we built mixed models as described above with those outcomes that were influenced by the intervention as dependent parameters, and with the adherence rates as independent parameters.

Results

Figure 7.1 shows the trial profile. We collected follow-up data for all units enrolled in the study, however, the number of the residents varied per measurement. In total, 883 residents provided informed consent and 793 residents were eligible and enrolled in the study with at least one measurement (T0-T5).



...
see next page

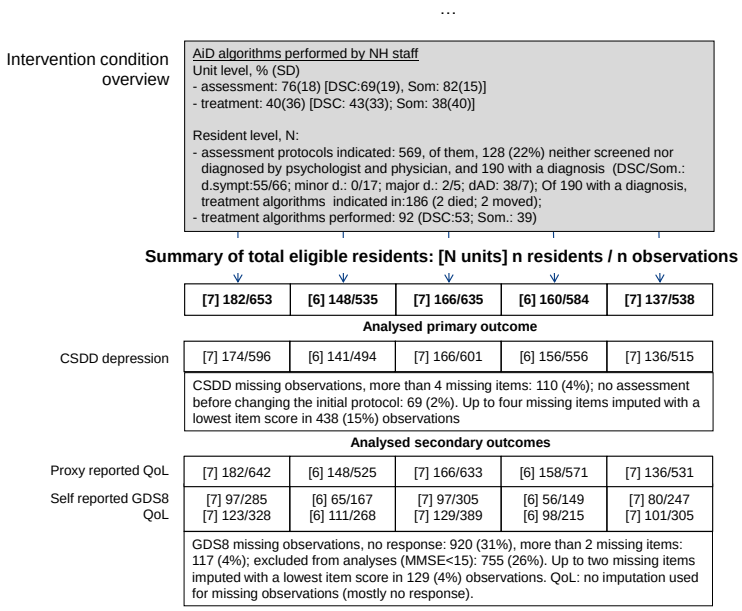


Figure 7.1 Study flow diagram

Page 112: Measurements in white represent usual care condition, and in grey - measurement in the intervention-condition; 1 unit included and randomized after T0.
Page 113: AiD algorithms performed by NH staff, unit level: the proportion of the residents for whom NH staff used the algorithms in relation to residents for whom the algorithms should be used;
Endpoint, primary outcome, depression based on the CSDD scores; Secondary outcomes, GDS8: depression outcomes based on reports of residents; QoL: quality of life reported by the clients (self-reported) and nursing staff (proxy-reported).

Abbreviations Page 112, N [n]: N is number of units, n is number of residents with informed consent; NH: nursing home; +IC: recruited and informed consent provided; -N(D): number residents deceased; -N(R): replaced to another unit; -N(W): IC withdrawal; T0–T5: measurements; res.: residents; Eligible: number of eligible residents; CSDD: numbers analyzed for the primary outcome measured using the Cornell Scale for Depression in Dementia; -N(ad): number of residents not assessed due to administrative reasons or logistic problems (e.g. resident not available during the measurement); -N(m): not included in primary analyses when more than 4 CSDD items were missing.
Abbreviations Page 113, DSC: dementia special care units; Som.: somatic units; d. sympt.: depressive symptoms (a screening instrument's score > cut-off) but no depression; minor d.: minor depression, 2 to 4 symptoms according to the DSM-IV-TR[39] criteria for major depression; major d.: major depression according to the DSM-IV-TR[39] criteria; dAD: depression according to the Provisional Diagnostic Criteria for Depression of Alzheimer's Disease (PDCdA)[40] in residents with dementia;
CSDD: Cornell Scale for Depression in Dementia; GDS8: Geriatric Depression Scale 8 items; QoL: quality of life assessed using visual analogue scale of the Euroqol-5D.

Intervention groups did not differ in unit size, number of residents included in the study, age, gender (Table 7.1), or in the eight outcome variables at unit's entry (proxy- and self-reported depression prevalence, severe depression prevalence, depressive symptoms and QoL). They differed in the number of units located in one of the four provinces (χ^2 [12, N=33]=23.3, *P*=0.026). For newly admitted residents at each T, we investigated whether groups differed in one of the eight outcome variables. The only difference in self-reported QoL at T2 (F(3, 6)=5.9, *P*=0.032) was considered being by chance at alpha=0.05 for 48 tests (six measurements x eight variables).

Table 7.1
Baseline cluster characteristics at unit's entry per intervention group (T0 for 32 units and T1 for one unit)

	Group 1	Group 2	Group 3	Group 4	Group 5		
Province, N (DSC)						Total	
- Noord-Brabant	4 (3)	4 (2)	1 (0)	4 (3)	4 (2)	17 (10)	
- Gelderland	0 (0)	0 (0)	0 (0)	2 (1)	2 (1)	4 (2)	
- Limburg	3 (1)	2 (1)	3 (1)	0 (0)	1 (0)	9 (3)	
- Zeeland	0 (0)	0 (0)	3 (1)	0 (0)	0 (0)	3 (1)	
Total N units (DSC)	7 (4)	6 (3)	7 (2)	6 (4)	7 (3)	33 (16)	
Characteristics, n						All groups	P Value
Unit size, M(SD)	28.1(15.1)	30.6(11.6)	26.7(5.7)	27.3(5.3)	25.4(7.4)	27.6(9.6)	0.921
n included, M(SD)	15.4(5.4)	16.0(8.0)	15.3(4.3)	16.8(6.8)	13.0(5.0)	15.2(5.2)	0.817
Age, M(SD)	84.1(1.1)	78.0(9.0)	81.4(2.4)	80.7(3.7)	78.0(10.0)	80.5(6.5)	0.393
Female, %(SD)	70.6(15.6)	67.0(16.0)	65.9(11.9)	64.8(9.8)	71.7(13.7)	68.2(12.2)	0.853

P values: ANOVA test, residents' characteristics at baseline were adjusted for clustering by comparing cluster means/proportions; M, SD, %: calculated at unit level.
Abbreviations, N: number of units; n included: number of residents enrolled in the study; DSC: dementia special care; M: mean; SD: standard deviation; %: proportion of residents in the unit;

Compared to somatic residents, DSC residents were older, had more cognitive impairments, scored higher on self-reported QoL, and were more often depressed according to proxies (Table 7.2). Those DSC residents who could be interviewed and had an MMSE >=15 (n=111 of 403) were less depressed according to GDS8. Dementia was diagnosed in 98% of the DSC and 13% of somatic residents (Table 7.3).

GDS8 correlated weakly with CSDD (df, 282; r, 0.23; *P*<0.001), and the same was found for self-reported QoL and proxy-reported QoL (df, 436; *r*, 0.10; *P*=0.029).

Table 7.2
Baseline characteristics of residents at entry
(first measurement after resident enrolment in the study)

	DSC		Somatic		Total		P
N residents	403		390		793		
Female, n (%)	280(69)		250(64)		530(67)		
Years of residence, M(SD)	1.7(2.2)		2.1(2.8)		1.9(2.5)		0.082
Adjusted for clustering†		n		n		n	
- Age*	83.4(1.5)	403	77.4(8.6)	390	80.3(6.9)	793	0.012
- MMSE	9.2(3.2)	293	20.2(3.3)	312	14.9(6.4)	605	<0.001
- Female*, %	70.1(13.5)	403	64.2(11.1)	390	67.1(12.5)	793	0.180
Endpoint, % M(SD)							
- CSDD depression	52.4(16.0)	382	40.6(12.0)	312	46.3(15.1)	694	0.022
Secondary outcomes, M(SD)							
- CSDD severe depression, %	22.8(12.0)	382	17.4(13.6)	312	10.0(12.9)	694	0.234
- CSDD score	8.5(1.9)	382	7.4(2.0)	312	8.0(2.0)	694	0.105
- GDS8 depression,%	18.5(24.4)	90	44.8(15.0)	240	32.0(23.9)	330	0.001
- GDS8 severe depression,%	10.9(16.8)	90	26.8(17.2)	240	19.1(18.6)	330	0.012
- GDS8 score	1.3(1.1)	90	2.6 (0.9)	240	1.9(1.2)	330	0.002
- QoL: proxy-reported	64.3(4.9)	399	61.0(5.8)	381	62.6(5.5)	780	0.079
- QoL: self-reported	70.4(7.4)	170	60.4(6.3)	279	65.3(8.4)	449	<0.001

Years of residence: length of stay in the unit from admission until first measurement resident enrolled;
Adjusted for clustering†: numbers are adjusted for clustering;
CSDD depression: proportion of residents with a CSDD score >7; CSDD severe depression: proportion of residents with a CSDD score > 11; GDS8 depression: proportion of residents with a GDS8 score >2; GDS8 severe depression: proportion of residents with a GDS8 score >4;
Significance P for difference between DSC and somatic units (t-test, residents' characteristics were adjusted for clustering by comparing cluster means/proportions).
Age*: can be considered comparable to 84 years in DSC and lower than 82 in somatic units in Dutch NHs; [51]
Female*: proportion can be considered lower than 77% in Dutch NHs.[51]
Abbreviations, n: number of residents, 16 DSC and 17 somatic units; DSC: dementia special care units; M: mean; SD: standard deviation; %: proportion; CSDD: the Cornell Scale for Depression in Dementia; GDS8: a shortened eight-item version of the Geriatric Depression Scale; QoL: quality of life visual analogue scale of the Euroqol-5D, reported by the clients (self-reported) and nursing staff (proxy-reported).

Table 7.3
Characteristics of residents, additional information

	DSC	Somatic	Total
Stepped-wedge observations in the trial, M(SD)			
- usual care (control) condition	2.0(1.5)	1.8(1.5)	1.9(1.5)
- intervention condition	1.8(1.6)	1.9(1.6)	1.8(1.6)
Mortality			
- n (%)	160(40)	109(28)	269(34)
- years residence in NH unit, M(SD)	2.7(2.5)	2.4(2.3)	2.6(2.4)
Proportion of residents in the trial (%)			
Married or partnered	34	29	32
Residing in the unit			
- single room	54	43	49
- two residents in room	36	48	42
- three residents in room	4	1	3
- four residents in room	7	7	7
Co-morbidity (%)			
- dementia diagnosis	98	13	56
- myocardial infarction	8	11	9
- congestive heart failure	12	21	17
- peripheral vascular disease	3	14	8
- cerebrovascular disease	30	56	42
- chronic pulmonary disease	13	23	18
- musculoskeletal and connective tissue disorders	6	8	7
- ulcer disease	3	8	5
- hemiplegia	7	34	20
- renal disease	11	18	14
- leukemia	0	0	0
- lymphoma	0	1	0
- diabetes mellitus	18	27	23
- malignant tumor	9	12	11

Stepped-wedge observations: number of 4-month measurements during the trial in the control and in the intervention conditions (range 0-5); years residence in NH unit: length of stay in the unit; Comorbidity data provided by NH physicians at T5 (N=605; missing data due to unavailable files in case of death or relocation)

Depression outcomes

Proxy-reported depression prevalence was 7.3 percentage points lower (95% Confidence interval [CI], -13.7 to -0.9; Cohen's d, -0.6) in the intervention condition in somatic units compared to usual care. No effect was found in DSC units, and no effect was found for severe depression in both DSC and somatic units. In the intervention condition, the CSDD score in somatic units was lower (-0.8; 95% CI, -1.4 to -0.1) than in usual care and not affected in DSC units (Table 7.4).

Self-reported depression outcomes measured with the GDS8 did not changed after introduction of AiD.

Explorative analyses showed that there was no significant influence of time factors (intervention and study duration) on the proxy-reported depression outcomes. The prevalence of severe GDS8 depression (-1.3%; 95% CI, -3.6% to 0.9%; reduced model *P*=0.049), and the GDS8 score (-0.1; 95% CI, -0.3 to 0.0; reduced model *P*=0.032) decreased with the time the resident was in the study.

When residents with low MMSE scores were included in the analyses, AiD had a significant effect on the GDS8 scores in somatic units (-0.2; 95% CI, -0.5 to 0.0; *P*=0.048).

Table 7.4
Effects of the AiD program (intention to treat analyses)

	DSC units		Somatic units		DSC/Somatic
Outcome	Effect (95% CI)	P	Effect (95% CI)	P	P
Endpoint					
- CSDD depression,% (ICC, 0.051)	0.6 (-5.6 to 6.8)	0.855	-7.3 (-13.7 to -0.9)	0.026	0.031
Secondary outcomes					
- severe CSDD depression,%	2.4 (-2.4 to 7.2)	0.331	-3.8 (-8.8 to 1.3)	0.141	0.031
- CSDD score	0.3 (-0.3 to 0.9)	0.379	-0.8 (-1.4 to -0.1)	0.018	0.004
- GDS8 depression*, %	-4.5 (-15.0 to 6.0)	0.405	-1.2 (-8.2 to 5.8)	0.733	0.554
- severe GDS8 depression*,%	-0.3 (-0.8 to 0.1)	0.167	-0.1 (-0.4 to 0.2)	0.425	0.396
- GDS8 score*	-0.3 (-0.7 to 0.1)	0.172	-0.1 (-0.4 to 0.2)	0.404	0.422
- QoL: self-reported**	3.4 (0.5 to 6.3)	0.023	3.4 (0.5 to 6.3)	0.023	0.366
- QoL: proxy-reported**	-2.0 (-3.5 to -0.5)	0.010	-2.0 (-3.5 to -0.5)	0.010	0.769

Models are adjusted for gender, age, region of country, time points, intervention effect and the interaction with the type of unit; DSC/Somatic: significance tested for the interaction with the type of unit, i.e. whether effect in DSC units differs from the effect in somatic units;
* both effect and interaction can be eliminated from a model without lowering the variance explained; data presented for a full model with effect and interaction;
** interaction effect with the type of unit was eliminated from a model without lowering the variance explained; data presented for the reduced model;
time factors (intervention and study duration) had no influence on the proxy-reported depression outcomes, and self-reported QoL; the time the resident was in the study was associated with the GDS8 score (slope, -0.1; 95% CI, -0.3 to 0.0); and improvement in proxy-reported QoL (slope, 1.4; 95% CI, 0.3 to 2.5).
Abbreviations, DSC: dementia special care; CSDD: Cornell Scale for Depression in Dementia; GDS8: shortened Geriatic Depression Scale with 8 items; CSDD depression: proportion of residents with a CSDD score >7; CSDD severe depression: proportion of residents with a CSDD score > 11; GDS8 depression: proportion of residents with a GDS8 score >2; GDS8 severe depression: proportion of

residents with a GDS8 score >4; QoL: quality of life visual analogue scale of the Euroqol-5D, reported by the clients (self-reported) and nursing staff (proxy-reported).

Quality of life

Self-reported QoL improved in the intervention condition (VAS points, 3.4; 95% CI, 0.5 to 6.3), with no difference between DSC and somatic units (Table 7.4).

Proxy-reported QoL decreased in the intervention compared to the usual care condition (-2.0; 95% CI, -3.5 to -0.5) with no difference between DSC and somatic units.

The time the resident was in the study was associated with improvement in self-reported QoL (1.4; 95% CI, 0.3 to 2.5; *P*=0.016) but there were no time-effects for proxy-reported QoL.

Table 7.5
Effects of the AiD program pathways. Intention to treat analyses adjusted for adherence rates.

	DSC units		Somatic units		DSC/Somatic
Outcome	Effect (95% CI)	P	Effect (95% CI)	P	P
Depression assessment procedures					
- CSDD depression, %	4.0 (-7.0 to 16.0)	0.457	-13.0 (-24.0 to -3.0)	0.015	0.017
- CSDD score	0.9 (-0.3 to 2.1)	0.147	-1.5 (-2.6 to -0.4)	0.007	0.002
- QoL: self-reported	8.2 (0.1 to 16.3)	0.047	5.0 (-0.3 to 10.2)	0.065	0.477
- QoL: proxy-reported	-5.0 (-8.4 to -1.6)	0.004	-5.6 (-8.6 to -2.5)	<0.001	0.768
Treatment protocols					
- CSDD depression, %	-5.0 (-19.0 to 10.0)	0.516	9.0 (-5.0 to 24.0)	0.219	0.184
- CSDD score	-0.8 (-2.2 to 0.7)	0.303	1.0 (-0.5 to 2.5)	0.199	0.102
- QoL: self-reported	-3.5 (-13.5 to 6.5)	0.490	-2.4 (-9.3 to 4.5)	0.491	0.857
- QoL: proxy-reported	4.1 (-0.1 to 8.3)	0.054	8.2 (3.9 to 12.4)	<0.001	0.186

Models are adjusted for gender, age, region of country, time points, adherence rates (extent to which NH teams used pathways for AiD depression assessment procedures and treatment protocols) and the interaction with the type of unit. DSC/Somatic: significance tested for the interaction with the type of unit, i.e. whether effect in DSC units differs from the effect in somatic units;
Abbreviations, DSC: dementia special care; CSDD: Cornell Scale for Depression in Dementia; GDS8: shortened Geriatic Depression Scale with 8 items; CSDD depression: proportion of residents with a CSDD score >7; QoL: quality of life visual analogue scale of the Euroqol-5D, reported by the clients (self-reported) and nursing staff (proxy-reported).

Program components

The DSC units had a lower adherence rate (Mean, 69%; SD, 19%) than somatic units (Mean, 82%; SD, 15%) for depression assessment, t(31)=2.1, *P*=0.045, while they did not differ in the adherence rates for treatment (Mean, 40%; SD, 36%). No difference in adherence rates was found between the five intervention groups. For all units, the adherence rate for depression assessment was higher than for treatment (paired t(31)=5.5, *P*<0.001).

Table 7.5 shows the effects when adjusted for the adherence rates. When depression assessment was performed, proxy-reported depression outcomes (significantly) and self-reported QoL (borderline significantly) improved, and proxy-reported QoL worsened in somatic units. In DSC units, proxy-reported QoL also worsened and self-reported QoL improved. Absolute effect-sizes were larger than for the whole program (endpoint Cohen's d, -1.1). When treatment algorithms were implemented, proxy-reported QoL showed a significant improvement in somatic units (8.2; 95% CI, 3.9 to 12.4; *P*<0.001), and borderline significant improvement in DSC units (DSC: 4.1, 95% -0.1 to 8.3; *P*=0.054).

Discussion

To the best of our knowledge, this study is the largest depression intervention trial performed in NHs to date and the first NH trial on a structural approach to depression management that includes assessment algorithms as well as pathways for collaborative treatment. The primary analysis showed that the AiD program can be used effectively for reducing depression prevalence and the severity of depressive symptoms in somatic units. The results of the secondary analyses on self-reported quality of life (QoL) supported the conclusion that AiD can be effective in somatic units and even suggested that a structural approach to depression management can be effective in DSC units. However, a negative effect of the whole program on proxy-reported QoL, and no significant effects on severe depression and on self-reported depression appear contradictory to the main findings.

Remarkably, improvement in self-reported QoL coincided with proxy-reported QoL worsening in the intervention condition. During the program introduction, caregivers were educated about depression and its consequences for the quality of life. They reported an increased awareness of and knowledge about depression.[19] This may have resulted in caregivers scoring residents with depression in the intervention condition as having lower health related QoL than before the intervention. Additional explorative analyses underpinned this assumption: the outcome worsened when assessment was implemented. More specifically, a model with individual assessment components (not shown) revealed worsening of the proxy-reported outcomes, but not the self-reported

outcomes, when a diagnostic procedure was implemented. Poor concordance between proxy- and self-reports on QoL[27] and depression scores[24] raises the fundamental question of who the most appropriate rater is when assessing subjective concepts in NH residents. The self-report should be the gold standard for the accuracy of QoL measures,[27] but whether and to what extent cognitive impairments affect the validity of self-reported QoL in NH residents is not clear. Regarding AiD, it is important to consider that the treatment adherence rate was associated with improvement of proxy-reported QoL, which implies that this outcome could be improved if more residents with depression would be treated.

We did not find an intervention effect using the GDS8. This may be, first, due to excluding residents with severe cognitive impairments from analyses,[28] which, along with other reasons for missing items, resulted in a loss of power (60% missing data). Since GDS8 was the dependent variable, imputing missing data under the missing at random assumption would lead to similar results as the mixed model analysis we used.[41] Re-analyzing our data including residents with low MMSE scores, we found an intervention effect in somatic residents. Although the validity of the GDS can be questioned in people with severe cognitive impairments, this finding suggests that an effect could be found in a larger sample. A poor concordance between the CSDD and GDS8 scores needs to be noted. In a recent study, Towsley and colleagues found a comparable poor concordance between proxy-reported and self-reported CSDD measurements in nursing homes.[24] They noted that although obtaining resident input when assessing depression is essential it is often clinically infeasible for implementation into NH routine practice. Therefore, caregivers should be educated in the most effective ways to assess depression. AiD program was introduced in this pragmatic trial through educating the nursing staff about depression and screening. Considering that the CSDD showed to be valid in our subsample,[42] the poor concordance between the CSDD and GDS8 scales may be explained by different aspects of the disorder they measure. In contrast to the CSDD, the GDS8 does not measure non-mood aspects of depression (e.g. cyclic functions and retardation). Those aspects might be affected by the intervention to a larger degree than GDS8 mood symptoms, which may explain why effects were found for proxy-reported depression only.

Although pragmatic research is concerned less than explanatory research with how and why an intervention works, additional analyses on adherence rates of AiD components provide useful data into the effectiveness of this complex intervention. Effect-sizes for outcomes on which the intervention as a whole had a significant effect were larger when assessment procedures were applied (Cohen's d, whole program: -0.6; assessment procedures: -1.1). The fact that treatment algorithms were not implemented optimally

may explain why using treatment pathways was not related to depression outcomes. Because the intervention effect may primarily reflect a non-specific effect induced by assessment and not the effect of collaborative treatment, residents with severe depression were unlikely to benefit in the trial. It remains unclear whether and how the residents with severe depression can benefit from a structural approach to depression management in NHs. Considering the primary effect could be attributed to assessment procedures, the lack of program effects on depression in DSC units may be explained by a lower extent to which assessment was conducted in DSC units than in somatic units. Another explanation for the lack of a significant effect in DSC units may lie in the convergence between depression and dementia regarding symptoms such as sleep disturbances, psychomotor retardation or agitation, and anhedonia.[43] The CSDD includes these and other symptoms, and it is possible that measured changes in depression were blurred by symptoms that are related to (the worsening of) dementia. As it can be expected, there were more residents with dementia in DSC units than in somatic units.

Comparison of the results of our study with other depression RCTs in NHs is complicated because most RCTs have studied a mono-therapy, in contrast to a multidisciplinary approach to depression management that includes structural assessment. Two depression RCTs were reported on integrated interventions combining psychiatric and nursing care (for a review see Collet et al.[12]). Brodaty et al.[44] compared a formula-driven psychogeriatric team case management with a consultative approach, and usual care in dementia. They also did not find significant effects in residents with dementia on the CSDD and 15-item GDS.[44] In another RCT, depression outcomes were improved by behavioral management in nursing and residential homes, training of caregivers by an old-age psychiatric hospital outreach team, and ongoing support to individual workers.[45] The training nursing staff received may have been an important contributing factor to the positive AiD results. More RCTs on multidisciplinary depression management are needed, especially in residents with dementia, and studies should be designed to determine the effect of caregiver training in addition to structural assessment and treatment.

Strengths of this trial include the use of the novel stepped-wedge cluster randomized design that allows an introduction of the intervention in all clusters before the end of the study. Other strengths are the large number of participating nursing home units and residents, the use of both self-reported and proxy-reported constructs for the outcomes, the use of evidence-based elements in the program, and applying no exclusion criteria for the residents, which may enlarge generalizability. However, generalizability of the results may be uncertain because of potential selection bias. Unfortunately, as we did not collect

data on residents without their consent, we could not compare our sample with non-participants. Depression prevalence in our sample may be considered within the range of prevalence reported in various NH studies using GDS versions, or the CSDD.[46-48] Other limitations of our study are worth noting. First, although NH staff reported that AiD is feasible and useful in a NH setting,[19] treatment pathways were not used for all depressed residents. Therefore, feasibility of the three treatment modules may be questioned. However, NH professionals could choose to work according to another method than prescribed by an AiD module because the focus of this RCT was not on specific treatment interventions but on a structural multidisciplinary approach to depression management. Although there was an indication that standardized depression assessment procedures resulted in positive depression and QoL outcomes, the low adherence to treatment pathways can be considered as disappointing, which calls for more research on the implementation of evidence-based depression treatment in NHs. Another limitation of our study concerns using a depression scale as an outcome instead of a comprehensive diagnostic procedure for determining depression. A comprehensive procedure was not possible in our study without breaking the NH professionals' blinding for the outcomes. Furthermore, to account for clustering in the units and for correlation over time, random effects models were used in our study. This means that the repeated measurements were modelled to have equal correlation between any two time points within a cluster. This is a conservative estimation of effects (larger standard errors). Using more specific correlation structures for the time points (e.g. autocorrelation) could provide more precision, but to the best of our knowledge, models for such structures have not been well developed.

To conclude, this study demonstrates that a structural approach to depression management in NHs could be effective, especially when standardized depression assessment procedures are used on a regular basis. More attention is needed for implementation of evidence-based treatment interventions in daily practice and on treatment of severe depression in nursing homes.

Literature

1. Fullerton CA, McGuire TG, Feng Z, Mor V, Grabowski DC. Trends in mental health admissions to nursing homes, 1999-2005. *Psych Serv* 2009;60(7):965-971.
2. Barca ML, Engedal K, Laks J, Selbaek G. A 12 months follow-up study of depression among nursing-home patients in Norway. *J Affect Disord* 2010;120(1-3):141-148.
3. Rovner BW, German PS, Brant LJ, Clark R, Burton L, Folstein MF. Depression and mortality in nursing homes. *JAMA* 27 1991;265(8):993-996.

4. Smalbrugge M, Pot AM, Jongenelis L, Gundy CM, Beekman AT, Eefsting JA. The impact of depression and anxiety on well being, disability and use of health care services in nursing home patients. *Int J Geriatr Psychiatr* 2006;21(4):325-332.
5. Bartels SJ, Horn SD, Smout RJ, et al. Agitation and depression in frail nursing home elderly patients with dementia: treatment characteristics and service use. *Am J Geriatr Psychiatry* 2003;11(2):231-238.
6. Dorland L, Pot A, Verbeek M, Depla M. In Dutch: [Psychische hulpverlening voor ouderen in verzorgings- en verpleeghuizen; deelstudie 7.] In *Monitor geestelijke gezondheidszorg ouderen, rapportage* 2006 Utrecht: Trimbos-Instituut; 2007.
7. Kramer D, Allgaier AK, Fejtkova S, Mergl R, Hegerl U. Depression in nursing homes: prevalence, recognition, and treatment. *Int J Psychiatry Med* 2009;39(4):345-358.
8. Damian J, Pastor-Barriuso R, Valderrama-Gama E. Descriptive epidemiology of undetected depression in institutionalized older people. *J Am Med Dir Assoc* 2010;11(5):312-319.
9. Oestergaard S, Moldrup C. Improving outcomes for patients with depression by enhancing antidepressant therapy with non-pharmacological interventions: a systematic review of reviews. *Public Health* 2011;125(6):357-367.
10. Levin CA, Wei W, Akincigil A, Lucas JA, Bilder S, Crystal S. Prevalence and treatment of diagnosed depression among elderly nursing home residents in Ohio. *J Am Med Dir Assoc* 2007;8(9):585-594.
11. Bharucha AJ, Dew MA, Miller MD, Borson S, Reynolds C, 3rd. Psychotherapy in long-term care: A review. *J Am Med Dir Assoc* 2006;7(9):568-580.
12. Collet J, de Vugt ME, Verhey FR, Schols JM. Efficacy of integrated interventions combining psychiatric care and nursing home care for nursing home residents: a review of the literature. *Int J Geriatr Psychiatr* 2010;25(1):3-13.
13. Psychiatry AGSaAAfG. Consensus statement on improving the quality of mental health care in U.S. nursing homes: management of depression and behavioral symptoms associated with dementia. *J Am Geriatr Soc* 2003;51(9):1287-1298.
14. Koopmans R, Zuidema S, Leontjevas R, Gerritsen D. Comprehensive assessment of depression and behavioral problems in long-term care. *Int Psychoger* 2010;22(07):1054-1062.
15. AGS/AAGP. The American Geriatrics Society and American Association for Geriatric Psychiatry recommendations for policies in support of quality mental health care in U.S. nursing homes. *J Am Geriatr Soc* 2003;51:1299-1304.
16. Koopmans RTCM, Lavrijsen J, Hoek JF, Went P, Schols JMGA. Dutch Elderly Care Physician: A new generation of nursing home physician specialists. *J Am Geriatr Soc* 2010;58(9):1807-1809.
17. Trimbos-instituut. In Dutch: [Addendum ouderen bij de multidisciplinaire richtlijn depressie]. Utrecht: Trimbos-instituut; 2008.
18. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390.
19. Leontjevas R, Gerritsen DL, Koopmans RTCM, Smalbrugge M, Vernooij-Dassen MJFJ. Process evaluation to explore internal and external validity of the Act in Case of Depression care program in nursing homes. *J Am Med Dir Assoc* 2012;13(5):488.e1-8.
20. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 20 2011;11(1):91.
21. Mdege ND, Man MS, Taylor Nee Brown CA, Torgerson DJ. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* 2011;64(9):936-948.

22. Bohlmeijer E, Steunenberg B, Leontjevas R, Mahler M, Daniel R, Gerritsen DL. *Dierbare herinneringen Protocol voor individuele life-review therapie*. Enschede: Universiteit Twente; 2010.
23. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23(3):271-284.
24. Towsley G, Neradilek MB, Snow AL, Ersek M. Evaluating the Cornell scale for depression in dementia as a proxy measure in nursing home residents with and without dementia. *Aging Ment Healt* 2012;16(7):892-901.
25. Chopra MP, Sullivan JR, Feldman Z, Landes RD, Beck C. Self-, collateral- and clinician assessment of depression in persons with cognitive impairment. *Aging Ment Health* 2008;12(6):675-683.
26. Achenbach TM, Krukowski RA, Dumenci L, Ivanova MY. Assessment of adult psychopathology: meta-analyses and implications of cross-informant correlations. *Psychol Bull* 2005;131(3):361-382.
27. Kane RL, Kane RA, Bershadsky B, et al. Proxy sources for information on nursing home residents' quality of life. *J Gerontol B Psychol Sci Soc Sci* 2005;60(6):S318-S325.
28. Jongenelis K, Gerritsen DL, Pot AM, et al. Construction and validation of a patient- and user-friendly nursing home version of the Geriatric Depression Scale. *Int J Geriatr Psychiatr* 2007;22(9):837-842.
29. Yesavage JA, Brinka TL, Roseb TL, et al. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* 1983;17(1):37-49.
30. Smalbrugge M, Jongenelis L, Pot AM, Beekman AT, Eefsting JA. Screening for depression and assessing change in severity of depression. Is the Geriatric Depression Scale (30-, 15- and 8-item versions) useful for both purposes in nursing home patients? *Aging Ment Health* 2008;12(2):244-248.
31. Group E. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16(3):199-208.
32. Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry* 1991;148(1):102-105.
33. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Use of the Cornell scale in nondemented patients. *J Am Geriatr Soc* 1988;36(3):230-236.
34. Leontjevas R, van Hooren S, Mulders A. The Montgomery-Asberg Depression Rating Scale and the Cornell Scale for Depression in Dementia: a validation study with patients exhibiting early-onset dementia. *Am J Geriatr Psychiatry* 2009;17(1):56-64.
35. Leontjevas R, Gerritsen DL, Vernooij-Dassen M, Smalbrugge M, Koopmans R. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents in dementia. *Am J Geriatr Psychiatry.* In press.
36. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;28(2):182-191.
37. Hannan P, Murray D. Gauss or Bernoulli? A Monte Carlo comparison of the performance of the linear mixed-model and the logistic mixed-model analyses in simulated community trials with a dichotomous outcome variable at the individual level. *Eval Rev.* 1996;20(3):338-352.
38. Verbeke G, Molenberghs G. *Linear Mixed Models for longitudinal data*. New York: Springer,; 2000.
39. APA. Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association; 2000.
40. Olin JT, Schneider LS, Katz IR, et al. Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002;10(2):125-128.

41. Molenberghs G, Kenwar MG. *Missing Data in Clinical Studies*. West Sussex: John Wiley & Sons, Lt; 2007.
42. Leontjevas R, Gerritsen DL, Vernooij-Dassen MJ, Smalbrugge M, Koopmans RT. Comparative Validation of Proxy-Based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in Nursing Home Residents With Dementia. *Am J Geriatr Psychiatry.* In press.
43. Kallenbach LE, Rigler SK. Identification and management of depression in nursing facility residents. *J Am Med Dir Assoc* 2006;7(7):448-455.
44. Brodaty H, Draper BM, Millar J, et al. Randomized controlled trial of different models of care for nursing home residents with dementia complicated by depression or psychosis. *J Clin Psychiat* 2003;64(1):63-72.
45. Proctor R, Burns A, Powell HS, et al. Behavioural management in nursing and residential homes: a randomised controlled trial. *Lancet* 1999;354(9172):26-29.
46. Snowdon J, Rosengren D, Daniel F, Suyasa M. Australia's use of the Cornell scale to screen for depression in nursing homes. *Australas J Ageing* 2011;30(1):33-36.
47. Smalbrugge M, Jongenelis L, Pot AM, Eefsting JA, Ribbe MW, Beekman AT. Incidence and outcome of depressive symptoms in nursing home patients in the Netherlands. *Am J Geriatr Psychiatry* 2006;14(12):1069-1076.
48. Mitchell AJ, Kakkadasam V. Ability of nurses to identify depression in primary care, secondary care and nursing homes-A meta-analysis of routine clinical accuracy. *Int J Nurs Stud* 2011;48(3):359-358.

Chapter

8

Activating strategies and psychotherapy reduce apathy and depressive motivational symptoms in nursing home residents

Ruslan Leontjevas, Steven Teerenstra, Martin Smalbrugge, Myrra J.F.J. Vernooij-Dassen, Ernst T. Bohlmeijer, Debby L. Gerritsen, Raymond T.C.M. Koopmans

Abstract

Background: The care program 'Act in case of Depression' (AiD) reduced depression prevalence in Dutch somatic nursing home (NH) units, but not in dementia special care (DSC) units. Apathy overlaps with depression in motivational symptoms but not in mood symptoms; apathy is common in dementia and can be influenced when treating depression.

Objective: To determine whether AiD influenced (1) apathy, (2) depressive motivational and mood symptoms, and (3) which AiD elements were effective.

Design, Setting and Participants: Stepped-wedge cluster-randomized trial between May 2009 and April 2011. 16 dementia special care (DSC) and 17 somatic NH units in four Dutch provinces were randomized in 5 intervention groups. 403 DSC and 390 somatic residents were recruited with no exclusion criteria except informed consent refusal.

Intervention: AiD prescribes pathways for structural assessment (a two-step screening procedure and a diagnostic procedure), multidisciplinary treatment (basic behavioral, psychotherapeutic, pharmacological), and monitoring of treatment effects. In the control condition, units provided usual care.

Measurements: Apathy, the 10-item version of the Apathy Evaluation Scale; motivational depressive symptoms, items of the Cornell Scale for Depression in Dementia (CSDD): retardation, loss of interest, appetite loss, weight loss, and lack of energy; mood symptoms, items: anxiety, sadness, irritability, agitation, diurnal variation, suicide, poor self-esteem, and pessimism of the CSDD.

Results: The whole AiD program reduced apathy in DSC units (-2.3; 95% CI, -3.3 to -1.3), and motivational symptoms in somatic units (-0.4; 95% CI, -0.7 to -0.1). The effect on apathy in DSC units was mainly attributable to behavioral activation strategies, whereas the effect on motivational symptoms in somatic residents was mainly attributable to psychotherapy. Apathy worsening was associated with pharmacological depression treatment. AiD did not influence mood symptoms.

Conclusions: The AiD program can be used for reducing apathy, and motivational depressive symptoms. Behavioral activation strategies in particular can be used effectively in DSC units, and psychotherapy in somatic residents. Mood symptoms are unlikely to be influenced by the program. More research is needed on improvement of depression and apathy management in NHs.

Introduction

A large two-year study on the effectiveness of the multidisciplinary depression care program Act in case of Depression (AiD)[1] was conducted in dementia special care (DSC) and somatic nursing home (NH) units in the Netherlands. Our analyses showed that AiD reduced depression prevalence and severity in somatic but not in DSC residents.[2] To better understand the effectiveness of AiD, its effect on apathy was analyzed in this study.

Apathy, which 'conventionally describes a lack of interest or emotion',[3] is traditionally considered a symptom of depression. Existing nomenclatures consider loss of interest or pleasure to be a principal symptom to diagnose depression, even when depressed mood is not present.[4,5] However, various studies support the concept that 'apathy is not depression'.[3,6] Apathy is increasingly recognized as a behavioral syndrome that can be discriminated from depression,[3,7,8] and is characterized by diminished motivation in combination with a lack of goal-directed behavior and goal-directed cognition, and a lack of emotional affect.[9-11] Nevertheless, although apathy can be seen as a separate syndrome, apathy and depression overlap. In several studies, the convergence between apathy and depression was attributable to so-called motivational symptoms including loss of interest, reactivity, psychomotor retardation, energy loss, and lack of insight, but not to mood-related symptoms such as sadness, feelings of guilt, and low self-esteem.[12,13]

The AiD care program is a complex intervention which includes a comprehensive depression assessment procedure, and pathways for psychosocial and pharmacological treatment of depression. Psychosocial treatment protocols consist of behavioral activation strategies, such as a pleasant-activities-plan,[14,15] and psychotherapy. Activating a client can also be beneficial for apathy.[16] On the contrary, pharmacological treatment of depression using antidepressants may induce apathy.[17,18] Because AiD treatment elements can influence apathy and depression differently, they can also influence overlapping motivational symptoms differently than mood symptoms. Therefore, while exploring the effect of AiD on apathy, it is important to focus on motivational and mood symptoms separately. Such an exploration is important not only for better understanding the effectiveness of AiD, but it can also contribute to the debate on apathy as a distinct syndrome.

Given that apathy, which is associated with quality of life,[19,20] may be affected by AiD treatment elements, it is important to monitor apathy when treating depression in NH residents. Monitoring apathy is especially indicated in dementia, because apathetic behavior is the most common non-cognitive disturbance in NH residents with dementia,[21-

[23] is associated with a faster cognitive and functional decline,[24] and with increased mortality in these residents.[25,26] Exploring the effects of the AiD elements on apathy can give insight into whether the AiD program should be fine-tuned preferably to reduce and not induce apathy.

Our first aim in this study was to determine the effect of AiD on apathy in both DSC and somatic units and to determine whether the effect was different in the two unit types. Our second aim was to explore the effects of AiD on depressive motivational and mood symptoms. Third, considering that AiD consists of several components, and different treatment elements can influence apathy differently from depression, we explored which AiD elements contributed to the results.

Methods

Design Overview

We designed a stepped wedge[27] trial with five intervention groups, six measurements, and the NH unit as the unit of randomization. Clusters crossed over from the control to the intervention condition at different time points, directly after a measurement each 4 months apart (Figure 6.2). The Medical Ethics Committee region Arnhem-Nijmegen rated the study and pronounced that it would not be burdensome for the participant.

We recruited units by convenience between February 2009 and May 2009. The Nijmegen University Network of NHs (UKON, www.uko-n.nl), a collaboration between 12 care organizations and the Department of Primary and Community Care of the Radboud University Nijmegen Medical Centre, invited its organizations to participate. In each NH site, no more than one DSC and one somatic unit were invited. Before T0, residents were recruited directly after unit inclusion. Newly admitted residents were recruited until the last measurement.

Setting and Participants

The residents were recruited from 16 dementia special care (DSC) and 17 somatic units by nursing staff members who did not have an impact on who was recruited; all residents or their legal representatives were to be approached. No exclusion criteria were used for residents with the exception of not providing informed consent. Each participating NH resident and/or the legal representative (in case the resident was incapable of giving informed consent) received written and verbal information prior to joining the AiD study.

Randomization and Interventions

Clusters were randomized to one of the 5 intervention groups by the researcher (RL), who was not involved in the recruitment, using computer-generated random numbers. If there were two units in one NH, they were randomized in pairs to avoid contamination bias. Recruited residents were assigned a unique code. The residents were informed about the study but not about the precise schedule of the unit cross over to the intervention condition nor about concrete program components used for them. Research staff administered the measurement instruments blinded to individual program components used for the resident.

Intervention: In short, the AiD care program consists of three components or phases for depression assessment, three treatment modules, and monitoring of the treatment results (see for more details Appendix A1). Each phase, with the exception of the first screening phase that was completed each 4th month in the AiD intervention condition, starts when indicated by the results of a preceding phase. The three assessment components are: a screening procedure by the nursing staff using a short observer-rated scale[28] (Phase 1); an extensive screening by the NH-unit psychologist using an interview-based instrument[29,30] (Phase 2), a diagnostic procedure by the unit psychologist and the elderly care physician (Phase 3). In Phase 4, the multidisciplinary team (psychologist, physician, and care team including nursing staff and recreational therapist) can provide up to three treatment modules: environmental and behavioral strategies including a pleasant activities plan and a day structure program (module 1) in case of depressive symptoms, i.e. a score higher than a cut-off in Phase 2; complementary to treatment module 1, psychotherapy in case of depression (module 2); and medication (module 3), especially if depression is severe. The psychologists and physicians may diverge from the content of individual treatment protocols if deemed necessary, but are requested to use the pathways for collaborative treatment. In the fifth phase of the AiD program, the treatment results are monitored by the psychologists and/or physician using the instruments from Phase 2 (see Appendix A1 for the program pathways).

The research team provided units in the intervention condition with program texts and practical tools, a 3.5 hour educational course about depression and AiD to the nursing staff, and a 3.5 hour training session to the psychologists about life review therapy.[21] A physician involved in the program development contacted the unit physician by phone to discuss a medication protocol.

No specific information about Aid was provided to NH staff and residents during the control condition. The units did not use any particular depression care program, and

depression was mostly assessed after indications of possible depression were raised by nursing staff, resident, or next of kin.[20]

Outcomes and Follow-up

Apathy was assessed using the abbreviated 10-item *Apathy Evaluation Scale (AES-10).* Lueken et al.[31] refined the original 18-item AES scale[32] for the NH population by eliminating items that either had lost specificity due to the mainly externally driven context in NHs, or were difficult to measure in residents with severe cognitive deficits. Each item of the AES-10 gives an example of apathetic behaviour. The answer categories are 1 = not at all characteristic, 2 = slightly characteristic, 3 = somewhat characteristic, and 4 = a lot characteristic, resulting in a scale ranging from 10 to 40. A higher total AES-10 score indicates more apathetic behaviour. The AES-10 was validated in Dutch NH residents with and without dementia against diagnostic criteria for apathy and was found to be a valid instrument for distinguishing apathetic from non-apathetic residents in a heterogeneous sample of residents with and without dementia.[33]

Depressive symptoms were assessed using the Cornell Scale for Depression in Dementia (CSDD).[29] The CSDD consists of 19 items, each rated as 0=absent, 1=mild or intermitted, and 2=severe. A higher total scale score indicates more severe depressive symptoms. The scale has been validated in patients with dementia,[29] without dementia,[34] as well as in NH residents with dementia.[35-37] To determine subscales, we performed factor analysis (eigenvalues greater than 1.0, varimax rotation, loadings of 0.4 and greater), and revealed factors which together explained 37% of the variance. Items of the first factor (20% of variance; items: retardation, loss of interest, appetite loss, weight loss, and lack of energy) constituted a *motivational* subscale. The sum of the items of the second factor (6% of variance; items: suicide, poor self-esteem, pessimism), fourth factor (3% of variance; items: irritability, agitation, diurnal variation), and fifth factor (2% of variance; items: anxiety, sadness) was used to compose a *mood* subscale score.

Adherence rate: To determine the effects of individual AiD elements, adherence rates were calculated per unit as the proportion of initiated cases of AiD components in relation to indicated cases in residents of the unit.

The CSDD and AES-10 were administered in an interview with the primary professional caregivers (credentials can be compared to those of a licensed practical nurse in the US) by the researcher, research assistant, and 32 graduate psychologists in their final year of MSc, none of whom were involved in providing the care program. Next to demographical data retrieved from the resident file, the standardized Mini-Mental-State Examination

(MMSE)[38] was administered in a structured interview with the resident for assessing global cognitive functioning. The interviewers were trained in administering the scales. In accordance with the initial protocol,[1] the CSDD should only be assessed in DSC residents. The protocol was changed after the first week of the T0 measurement: The CSDD was administered in both DSC and somatic units, which makes a comparison between the two unit types possible using the same instrument.

Statistical analyses

We performed basic analyses using SPSS 17.0.0 (Chicago, IL). To compare baseline characteristics and adherence rates of AiD components between intervention groups, we used ANOVA on the unit means for continuous variables and proportions based on dichotomous variables at resident level. For newly admitted residents, we investigated at each T whether groups differed for the 3 outcome variables: apathy, motivational and mood symptoms. A t-test was used to compare DSC and somatic units. Chi-square tests were used for categorical variables of units. For a maximum of 4 missing individual CSDD items, a 0 score was imputed.[39] The same conservative imputation (lowest score = 1) was performed for a maximum of two missing AES-10 items.

To account for the nesting of measurements within subjects, and subjects within units, continuous outcomes were fitted using linear mixed models with random effects for units and for subjects nested within units (SAS software 9.2, SAS Institute Inc., North Carolina). Primary dependent outcomes were the AES-10, and two CSDD subscale scores. All estimates (here and below) were adjusted for age, gender, time trends (T0-T5), region of the country (province) and type of unit (somatic or DSC). The main effect of the intervention and its interaction with type of unit were tested using likelihood ratio tests comparing a model with those factors to, first, a model without the interaction and, second, to a model without the main effect and interaction.[40] Estimates of the main effect and possibly the interaction were taken from the most reduced model of which the fit was not significantly worse. Cohen's d is determined as the intervention effect divided by a standard deviation Exploratively, the influence of the duration of the intervention was investigated by first building a model with additional linear and quadratic terms for the duration and their interaction terms with the type of unit, and then reducing it to the smallest model of which the fit was not significantly worse. On the unit level, the duration of the study was conceptualized by the number of four-month periods (the T measurements were held approximately each four months) in the intervention condition. On the resident level, the number of four-month periods the resident participated in the

study was used for the time influence, and the number of periods the resident was in the unit in the intervention condition was used for the duration of the intervention.

A probability value of less than 0.017 with Bonferroni correction for three main variables (0.05/3) was considered significant.

To assess the influence of individual AiD phases, a model with the adherence rates of each phase, their interaction effect with the type of unit and duration factors that were of influence by the previous analyses was compared to the most reduced model that did not have a significantly worse fit.

Control condition

No specific information about AiD was provided to NH staff and residents during the control condition. In the control condition, the units did not use any structural approach to depression management: depression was assessed after indications of possible depression were raised by nursing staff, resident, or next of kin,[20] and teams did not use multidisciplinary pathways for depression treatment.

Results

Figure 8.1 shows the trial profile. Before T0, IC was provided by 547 (62%) residents or their representatives. We collected follow-up data for all units enrolled in the study, however, the number of residents varied per measurement. In total, 403 (51%) DSC and 390 (49%) somatic residents were involved in the trial. Intervention groups did not differ in unit size (Mean, 25.4; SD, 9.6), the number of residents included in the study (Mean, 15.2; SD, 5.2), age (Mean, 80.5; SD, 6.5) and female gender (68.2%; SD, 12.2%) of the residents enrolled (see Chapter 7). The groups differed significantly in the number of units located in one of the four provinces (Chi-square χ^2 [12, N=33]=23.3, *P*=0.026). The groups did not differ in the outcome measures at unit's entry. The difference in mood symptoms at T2 (F[4, 16]=4.2; *P*=0.016) was considered being by chance for 18 tests (3 variables, 6 time points) at alpha=0.05.

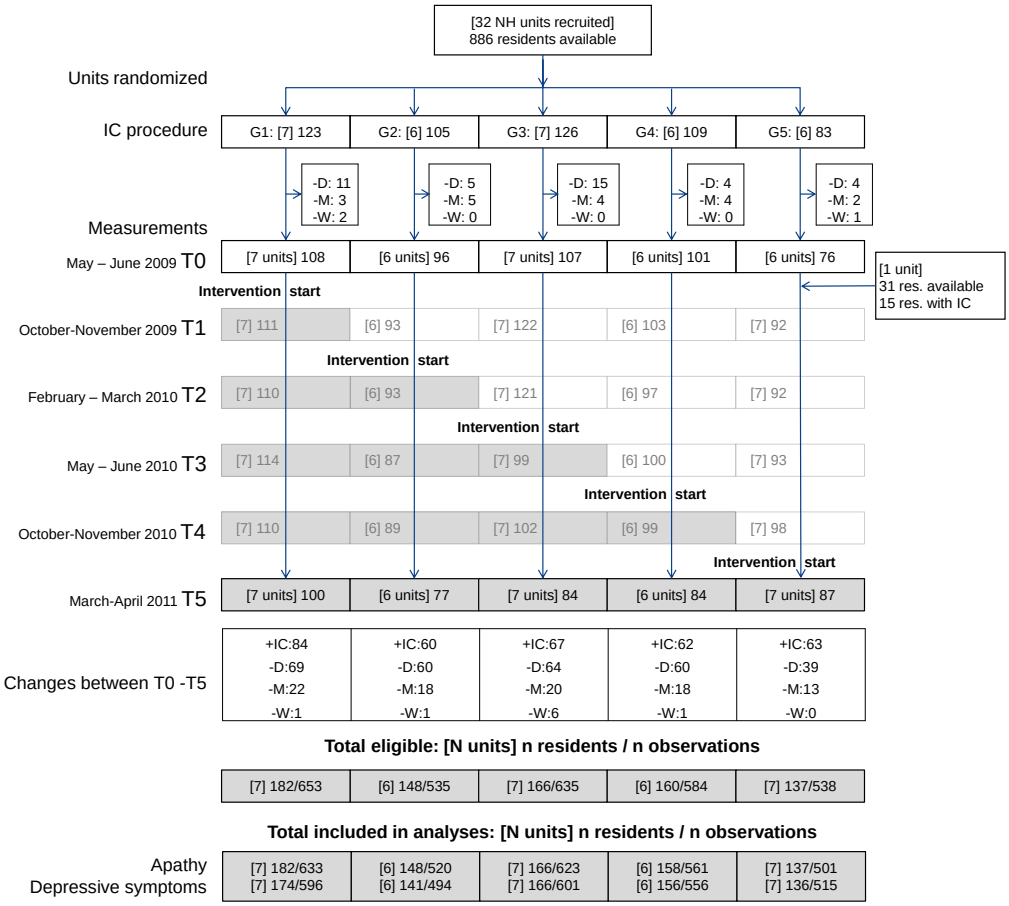


Figure 8.1 Study flow diagram

Measurements: clusters crossed over from the control to the intervention condition at different time points, directly after a measurement each 4 months apart. Consequently, the last follow up at T5 in the intervention condition varies from 4 months to approximately 20 months.
Apathy: the 10-item version of the Apathy Evaluation Scale.
Missing values were due to problems with scoring by proxies (all scales), and due to 69 residents (all groups represented) that were not assessed for the CSDD during the first week of the T0 measurement, which was according to the former protocol prescribing the use of the CSDD in DSC residents only (see Methods and figure 7.1).
Abbreviations, NH: nursing home; G=group; [N] n: N, number of units; n, number of residents with informed consent (IC); IC: informed concent; -D: deceased; -M: replaced to another unit, moved; -W: IC withdrawal; T0 – T5: measurements; +IC: recruited after T0; 1 unit included: paired with a previously randomized unit from the same NH to avoid contamination; res.: residents; observations: residents with available measurements.

At inclusion of the residents in the study, DSC residents were older, had more cognitive impairments, and more severe apathy and mood symptoms than somatic residents. They did not differ in gender, and motivational symptoms (Table 8.1).

Table 8.1
Residents' characteristics at entry

	DSC	Somatic	*P*	Total
Units, N (N DSC)	16	17		33
Residents, n (%)	403 (51)	390 (49)		793
Female, % (SD)	62.7 (36.7)	65.0 (34.0)	0.702	66.9 (25.9)
Age	83.6 (3.6)	76.9 (10.9)	<0.001	80.5 (7.5)
Mean (SD) /n				
- Cognition (MMSE)	10.6 (4.6) /293	19.5 (5.5) /312	<0.001	14.5 (6.5) /605
- Apathy (AES-10)	26.6 (6.6) /399	22.7 (7.7) /385	0.001	25.2 (5.2) /784
- Depressive symptoms				
Motivational	2.0 (2.0) /382	2.0 (1.0) /312	0.849	1.9 (1.9) /694
Mood	3.7 (2.7) /382	2.5 (1.5) /312	<0.001	3.0 (2.0) /694

P values: t-test for DSC versus Somatic, residents' characteristics at baseline were adjusted for clustering by comparing cluster means/proportions.
Cognition measured with the MMSE (minimum: 0, maximum: 30); Apathy, AES-10 (minimum:10; maximum: 40); Motivational symptoms, sum of CSDD-items: retardation, loss of interest, appetite loss, weight loss, and lack of energy (minimum: 0; maximum 10); Mood symptoms, sum of CSDD-items: anxiety, sadness, irritability, agitation, diurnal variation, suicide, poor self-esteem, pessimism (minimum: 0; maximum: 16).
Abbreviations, DSC: dementia special care units; MMSE: Mini-Mental-State Examination; AES-10: Apathy Evaluation Scale, 10 items; CSDD: Cornell Scale for Depression in Dementia.

DSC units showed a lower combined adherence rate for depression assessment (recognition, screening and diagnosing) than somatic units. No differences were found for individual phases. Intervention groups did not differ in adherence rates (Table 8.2).

Table 8.2
Extent to which NH teams used AiD pathways: proportion (%) of the residents receiving the AiD components in relation to residents who should receive the phase or treatment module, Mean % (SD) [N units]

	DSC	Somatic	Total	*P* type of unit	*P* groups
AiD phases					
Depression assessment					
- Phase 1 Detection	82 (17) [16]	89 (14) [17]	86 (16) [33]	0.201	0.292
- Phase 2 Screening	47 (39) [16]	64 (32) [17]	55 (36) [33]	0.183	0.236
- Phase 3 Diagnosing	52 (39) [15]	53 (43) [16]	52 (41) [31]	0.926	0.279
Total depression assessment	69 (19) [16]	82 (15) [17]	76 (18) [33]	0.045	0.394
Treatment					
- Phase 4 module 1	48 (42) [16]	35 (38) [16]	42 (40) [32]	0.337	0.967
- Phase 4 module 2	17 (35) [14]	49 (48) [15]	33 (45) [29]	0.056	0.607
- Phase 4 module 3	44 (45) [15]	37 (48) [15]	40 (46) [30]	0.666	0.485
Total treatment	43 (33) [16]	38 (40) [16]	40 (36) [32]	0.745	0.729
Phase 5 Monitoring	22 (32) [15]	14 (27) [16]	18 (30) [31]	0.452	0.548

Significance, *P* type of unit: significance tested for the difference between DSC and somatic units using t-test;
P groups: significance tested for the difference between groups using ANOVA-test;
Depression assessment procedures contain phases 1, 2 and 3; Treatment contains three modules, module 1: day program and pleasant activities plan; module 2: psychotherapy by psychologist; module 3: (re-) considering medication; difference between Depression assessment and Treatment adherence rates: Mean difference (SD), 0.35 (0.35), paired t(31)=5.5, *P*<0.001
Detection: screening by nursing staff; Screening: screening by psychologist; Diagnosing: diagnostic procedure by psychologist and physician; Treatment module 1: behavioral activation strategies; module 2: psychotherapy or mediative therapy; module 3: pharmacological treatment.
Abbreviations, [N units]: number of units with residents for whom an AiD component should be performed (missing values are due to not indicated components); DSC: dementia special care units

Apathy

Table 8.3 shows the effects of the AiD program as compared to usual care. Intention-to-treat analyses showed a reduction of apathy by 2.3 AES points in the DSC units (95% CI, -3.3 to -1.3; *P*<0.001; Cohen's d, -0.35), and no significant change of apathy in somatic units. This difference in the effect between the DSC and somatic units was significant (difference in effect, -2.8; 95% CI, -4.2 to -1.5; *P*<0.001). The time the resident participated in the study was associated with an increase of the AES scores in the DSC units. The time the resident

was in the intervention condition was associated with an AES increase in both DSC and somatic units.

Of the AiD elements, the extent to which treatment module 1 (pleasant activities plan and day structure program) was used in the units was significantly associated with apathy decrease in the DSC units (slope, -4.8; 95% CI, -6.7 to -2.9; *P*<0.001; Cohen's d, -0.73). In somatic units, module 2 (psychotherapy) showed apathy decrease (slope, -3.2; 95% CI, -5.1 to -1.3; *P*<0.001; Cohen's d, -0.42). Adherence rate of phase 2 screening by psychologist in somatic units, module 2 in DSC units (borderline significance), and module 3 pharmacological treatment in both DSC and somatic units were associated with an AES increase.

Depressive symptoms

AiD had a significant effect on motivational symptoms in somatic units (effect estimate, -0.4; 95% CI, -0.7 to -0.1; *P*=0.008; Cohen's d, -0.40), but not in DSC units. Mood symptoms were not influenced by the intervention. The time factors could be deleted from the models.

Of the AiD elements, the adherence rate of phase 2 screening was negatively associated with motivational symptoms, while phase 3 diagnosing - positively. Module 1 was associated with less motivational symptoms in DSC units (slope, -0.9; 95% CI, -1.3 to -0.4; *P*<0.001; Cohen's d, -0.45) and module 2 in somatic units (slope, -0.8; 95% CI, -1.4 to -0.3; *P*=0.002; Cohen's d, -0.80). Module 2 was associated with higher scores in DSC units (slope, 1.3; 95% CI, 0.6 to 2.0; *P*<0.001; Cohen's d, 0.65).

Table 8.3
Effects of the AiD program and its components on apathy, and on depressive motivational and mood symptoms.

	DSC units		Somatic units		DSC vs Somatic
	Effect (95% CI)	P value	Effect (95% CI)	P value	P value
Apathy					
Intervention effect	-2.3 (-3.3 to -1.3)	<0.001	0.5 (-0.5 to 1.6)	0.300	<0.001
Intervention duration*	0.7 (0.3 to 1.1)	<0.001	0.7 (0.3 to 1.1)	<0.001	*
Study duration	0.9 (0.5 to 1.4)	<0.001	0.3 (-0.2 to 0.7)	0.266	<0.001
Individual AiD components reduced model					
- Phase 2 Screening	-0.9 (-2.5 to 0.7)	0.271	2.5 (0.4 to 4.6)	0.019	0.012
- Treatment module 1	-4.8 (-6.7 to -2.9)	<0.001	1.4 (-1.4 to 4.3)	0.320	<0.001
- Treatment module 2	2.5 (-0.1 to 5.1)	0.061	-3.2 (-5.1 to -1.3)	0.001	0.001
- Treatment module 3*	1.8 (0.5 to 3.2)	0.009	1.8 (0.5 to 3.2)	0.009	*
Motivational symptoms					
Intervention effect	0.0 (-0.2 to 0.3)	0.854	-0.4 (-0.7 to -0.1)	0.008	0.011
Individual AiD components reduced model					
- Phase 2 Screening	-0.8 (-1.5 to -0.2)	0.015	-0.8 (-1.5 to -0.2)	0.015	*
- Phase 3 Diagnosing	1.3 (0.6 to 1.9)	<0.001	1.3 (0.6 to 1.9)	<0.001	*
- Treatment module 1	-0.9 (-1.3 to -0.4)	<0.001	0.2 (-0.6 to 1.0)	0.572	0.012
- Treatment module 2	1.3 (0.6 to 2.0)	0.001	-0.8 (-1.4 to -0.3)	0.002	<0.001
Mood symptoms					
Intervention effect**	0.2 (-0.2 to 0.5)	0.385	-0.3 (-0.6 to 0.1)	0.100	**

Models are adjusted for gender, age, region of country, time points, AiD phases' adherence rates (extent to which NH teams used the phase) and the interaction with the type of unit; DSC/Somatic: significance tested for the interaction with the type of unit, i.e. whether effect in DSC units differs from the effect in somatic units.
Apathy: AES-10 (minimum:10; maximum: 40); Motivational symptoms, sum of CSDD-items: retardation, loss of interest, appetite loss, weight loss, and lack of energy (minimum: 0; maximum 10); Mood symptoms, sum of CSDD-items: anxiety, sadness, irritability, agitation, diurnal variation, suicide, poor self-esteem, pessimism (minimum: 0; maximum: 16).
Intervention effect: effect of the whole program; Intervention duration: effect of the time the resident was in the intervention condition (number of 4-month periods); Study duration: the number of 4 month periods the resident was in the study.
Individual AiD components reduced model: a model that still has a non-significantly worse fit than a full model with all AiD components, estimates were taken from the most reduced model of which the fit was not significantly worse, other AiD components were deleted from a model without significant influence on the variance explained; Screening: screening by psychologist; Diagnosing: diagnostic procedure by psychologist and physician; Treatment module 1: behavioral activation strategies; module 2: psychotherapy or mediative therapy; module 3: pharmacological treatment.
*: interaction effect with the type of unit could be eliminated from a model;
**: both effect and interaction can be eliminated from a model.
Abbreviations, DSC: dementia special care units; AES-10: Apathy Evaluation Scale, 10 items; CSDD: Cornell Scale for Depression in Dementia.

Discussion

Our analyses showed that the whole care program 'Act in case of Depression' reduced apathy in DSC but not in somatic units. In contrast, the depressive motivational symptoms were reduced in somatic but not in DSC units, which was in agreement with the intervention effect on depression outcomes presented in chapter 7. Furthermore, analyses showed that depressive mood symptoms were not reduced.

Revealed effects of individual program components provided more insight into the effectiveness of the complex intervention, and into different effects of AiD on apathy and depressive symptoms. Module 1 in DSC units was associated with a decline in both apathy and depressive motivational symptoms. This can be explained by behavioral activation strategies used in this module.[16] On the contrary, there was an indication that module 2 in DSC units was associated with higher apathy and motivational symptoms scores. This module included mediative therapy in dementia by psychologists and it was implemented poorly in DSC residents (17% of the indicated cases). It is possible that the psychologists implemented the module mainly in those residents who showed worse symptomatology. In contrast to the DSC units, the adherence rate for module 2 was more acceptable (49%) in somatic units, and it was associated with reduction of apathy and motivational symptoms. Life-review therapy was the first choice for residents without severe cognitive impairments, and the effect of this therapy has been reported regarding depression in different studies.[41] Hsieh et al. reported an apathy decrease due to group reminiscence therapy, which is related to life-review therapy, in NH residents with mild to moderate dementia.[42] More research on the effect of individual or group reminiscence therapy on apathy is welcome in both somatic and DSC residents.

In contrast to psychosocial strategies, which showed congruent effects on apathy and motivational symptoms, pharmacological treatment was not related to motivational symptoms whereas apathy increased when module 3 was implemented. As mentioned, the use of antidepressants can induce apathy.[17,18] Yet, we did not adjust our analyses for the actual use of medication and it is possible that physicians started to use the AiD medication protocol more often in those residents who showed more apathetic behavior.

Next to this different result for module 3 on apathy and motivational symptoms, depression assessment phases showed different associations with apathy and motivational symptoms too. Screening by the psychologist was associated with reduced apathy in DSC units, and also with reduced motivational symptoms in DSC and somatic units. This suggests a non-specific effect above treatment modules. However, screening

by the psychologist was associated with more apathy in somatic units. Phase 2 was initiated only in 55% of prescribed cases and it is possible that the psychologists decided to specifically screen those somatic residents in whom caregivers reported an increase in apathetic symptoms. In this regard it is interesting that the higher adherence rate for diagnosing was associated with increased motivational symptoms but not with apathy. It is probable that psychologists were able to differentiate apathy from depressive symptoms using screening instruments in phase 2 and proceeded, in accordance with the AiD program, with diagnosing of those residents who showed depression and not only apathy. However, another explanation for the higher motivational scores after diagnosing is also possible. This may be due to caregivers' increased awareness of and knowledge about depression.[43] Caregivers, who were interviewed for depression scales, may rate the residents as more 'sick' after an official depression diagnosis was made. Such scoring bias after diagnosing would not necessarily affect apathy scores, which corresponded with our finding. Future studies are needed to address a possible scoring bias by proxies after diagnosing depressed residents.

Unfortunately, there was an indication that apathy worsened with time in DSC units independently of the intervention, and the intervention effect on apathy declined over time in both DSC and somatic units. The former may point to worsening of dementia.[44] The latter may imply that NH staff encountered problems in maintaining psychosocial treatment strategies that were effective for apathy. Another reason may reflect the use of antidepressants. Considering that module 3 and the effect drop over time were associated with apathy but not with depressive (motivational nor mood) symptoms, the negative effect of module 3 on apathy could be cumulative over time. Indeed, more research is needed on a longitudinal effect of antidepressants on apathy.

In sum, in DSC residents, the effect on apathy was mainly attributable to activating strategies (module 1) showing an effect-size of almost two thirds of the standard deviation. Although to a lesser degree than for apathy, module 1 was also effective for motivational symptoms in these residents. This effect, however, was overshadowed by the worsening of motivational symptoms due to causes that were associated with more diagnosing and poorly implemented mediative therapy. The activating strategies were not effective in somatic residents, which could be due to somatic disabilities impeding activities in this population. On the contrary, psychotherapy in these residents was the main contributor to reducing motivational symptoms. Psychotherapy was also effective for apathy in somatic residents, but its effect was overshadowed by apathy worsening due to medication or non-specific effects. To conclude, behavioral strategies of module 1 in

DSC residents were more effective for apathy than for depressive motivational symptoms, whereas psychotherapy in somatic residents was more effective for depressive motivational symptoms than for apathy. More research is needed on how psychotherapy can be used beyond activating strategies in DSC residents, and how activating strategies can be improved for somatic residents.

Several limitations of our study are worth noting. First, the care program was not developed initially for the recognition and treatment of apathy. Future research is needed into whether a comprehensive care program aimed at apathy is more effective than AiD. Second, generalizability of our results may be questioned due to potential selection bias. We did not expect that 40% to 50% of the residents in the units would not provide informed consent in the study.[43] Depression prevalence in our sample was comparable to that in other studies.[2] Apathy prevalence could not be compared with other figures as the AES-10 was not used in epidemiological studies on apathy yet. Another study limitation has been mentioned earlier: we did not account for the use of antidepressants. A final question may be whether the four month period between the measurements was sufficient to reach effects on outcomes. However, only one of five intervention groups had this short intervention period, so the effect of this is likely to be small.

Conclusion

This study shed more light on previously reported results of the effectiveness of the AiD care program. To reduce both apathy and motivational symptoms, AiD activating strategies can be used in DSC units, and psychotherapy in somatic units. This is an important finding for the improvement of NH care. Our findings can also be important for the ongoing discussion on the concept of apathy. Although some AiD treatment elements showed congruent effects on apathy and motivational symptoms, there were also indications for the concept that 'apathy is not depression':[6] apathy, but neither mood nor motivational depressive symptoms, worsened in the units with a higher adherence rate for pharmacological depression treatment.

Literature

1. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011;11:91.
2. Leontjevas R, Gerritsen DL, Smalbrugge M, et al. Effects of the nursing home multidisciplinary care program 'Act in case of Depression' including structural adepression assessment and treatment pathways: A stepped wedge cluster-randomized trial. Submitted.
3. Ishii S, Weintraub N, Mervis JR. Apathy: a common psychiatric syndrome in the elderly. *J Am Med Dir Assoc* 2009;10:381-393.
4. APA. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association, 2000.
5. Olin JT, Schneider LS, Katz IR, et al. Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002;10:125-128.
6. Levy ML, Cummings JL, Fairbanks LA, et al. Apathy is not depression. *J Neuropsych Clin N* 1998;10:314-319.
7. Starkstein SE, Leentjens AF. The nosological position of apathy in clinical practice. *J Neurol Neurosurg Psychiatry* 2008;79:1088-1092.
8. Van Reekum R, Stuss DT, Ostrander L. Apathy: why care? *J Neuropsych Clin N* 2005;17:7-19.
9. Marin RS. Apathy: a neuropsychiatric syndrome. *J Neuropsych Clin N* 1991;3:243-254.
10. Starkstein SE, Petracca G, Chemerinski E, et al. Syndromic validity of apathy in Alzheimer's disease. *Am J Psychiatry* 2001;158:872-877.
11. Robert P, Onyike CU, Leentjens AF, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatry* 2009;24:98-104.
12. Ready RE, Ott BR, Grace J, et al. Apathy and executive dysfunction in mild cognitive impairment and Alzheimer disease. *Am J Geriatr Psychiatry* 2003;11:222-228.
13. Leontjevas R, van Hooren S, Waterink W, et al. Apathy and depressive mood in nursing home patients with early-onset dementia. *Am J Alzheimer's dis and other dem* 2009;24:341-348.
14. Teri L, Logsdon RG, Uomoto J, et al. Behavioral treatment of depression in dementia patients: a controlled clinical trial. *J Gerontol B Psychol Sci Soc Sci* 1997;52:P159-166.
15. Verkaik R, Francke AL, van Meijel B, et al. The effects of a nursing guideline on depression in psychogeriatric nursing home residents with dementia. *Int J Geriatr Psychiatr* 2011;26(7):723-732.
16. Brodaty H, Burns K. Nonpharmacological Management of Apathy in Dementia: A Systematic Review. *Am J Geriatr Psychiatry* 2012;20(7):549-564.
17. Barnhart WJ, Makela EH, Latocha MJ. SSRI-induced apathy syndrome: a clinical review. *J Psychiatr Pract* 2004;10:196-199.
18. Settle EC, Jr. Antidepressant drugs: disturbing and potentially dangerous adverse effects. *The Journal of clinical psychiatry* 1998;59(16):25-30.
19. Samus QM, Rosenblatt A, Steele C, et al. The association of neuropsychiatric symptoms and environment with quality of life in assisted living residents with dementia. *The Gerontologist* 2005;45(1):19-26.
20. Gerritsen DL, Jongenelis K, Steverink N, et al. Down and drowsy? Do apathetic nursing home residents experience low quality of life? *Aging Ment Health* 2005;9:135-141.
21. Aalten P, de Vugt ME, Jaspers N, et al. The course of neuropsychiatric symptoms in dementia. Part I: findings from the two-year longitudinal Maasbed study. *Int J Geriatr Psychiatry* 2005;20:523-530.

22. Zuidema SU, Derksen E, Verhey FR, et al. Prevalence of neuropsychiatric symptoms in a large sample of Dutch nursing home patients with dementia. *Int J Geriatr Psychiatry* 2007;22:632-638.
23. Lyketsos CG, Olin J. Depression in Alzheimer's disease: overview and treatment. *Biol Psychiatry* 2002;52:243-252.
24. Starkstein SE, Jorge R, Mizrahi R, et al. A prospective longitudinal study of apathy in Alzheimer's disease. *J Neurol Neurosurg Psychiatry* 2006;77:8-11.
25. Diesfeldt HF, van Houte LR, Moerkens RM. Duration of survival in senile dementia. *Acta Psychiat Scand* 1986;73:366-371.
26. Van Dijk PT, Dippel DW, Habbema JD. A behavioral rating scale as a predictor for survival of demented nursing home patients. *Arch Gerontol Geriatr* 1994;18:101-113.
27. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;28:182-191.
28. Leontjevas R, Gerritsen DL, Vernooij-Dassen MJFJ, et al. The Nijmegen Observer-Rated Depression Scale for Detection of Depression in Nursing Home Residents. *Int J Geriatr Psychiatr* In press.
29. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23:271-284.
30. Yesavage JA, Brinka TL, Roseb TL, et al. Development and validation of a geriatric depression screening scale: A preliminary report. *J Psychiatr Res* 1983;17:37-49.
31. Lueken U, Seidl U, Volker L, et al. Development of a short version of the Apathy Evaluation Scale specifically adapted for demented nursing home residents. *Am J Geriatr Psychiatry* 2007;15:376-385.
32. Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Res* 1991;38:143-162.
33. Leontjevas R, Evers-Stephan A, Smalbrugge M, et al. A comparative validation of the abbreviated Apathy Evaluation Scale (AES-10) with the Neuropsychiatric Inventory apathy subscale against diagnostic criteria of apathy. *J Am Med Dir Assoc* 2011;13(3):308.e1-6.
34. Alexopoulos GS, Abrams RC, Young RC, et al. Use of the Cornell scale in nondemented patients. *J Am Geriatr Soc* 1988;36:230-236.
35. Leontjevas R, Gerritsen DL, Vernooij-Dassen M, et al. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents in dementia. *Am J Geriatr Psychiatry* In Press.
36. Leontjevas R, van Hooren S, Mulders A. The Montgomery-Asberg Depression Rating Scale and the Cornell Scale for Depression in Dementia: a validation study with patients exhibiting early-onset dementia. *Am J Geriatr Psychiatry* 2009;17:56-64.
37. Barca ML, Engedal K, Selbaek G. A reliability and validity study of the Cornell Scale among elderly inpatients, using various clinical criteria. *Dement Geriatr Cogn Disord* 2010;29:438-447.
38. Molloy DW, Alemayehu E, Roberts R. Reliability of a Standardized Mini-Mental State Examination compared with the traditional Mini-Mental State Examination. *Am J Psychiatry* 1991;148:102-105.
39. Leontjevas R, Gerritsen DL, Vernooij-Dassen M, et al. Comparative validation of proxy-based Montgomery-Asberg Depression Rating Scale and Cornell Scale for Depression in Dementia in nursing home residents in dementia. *Am J Geriatr Psychiatry* In Press.
40. Verbeke G, Molenberghs G. *Linear Mixed Models for longitudinal data*. New York: Springer, 2000.
41. Bohlmeijer E, Smit F, Cuijpers P. Effects of reminiscence and life review on late-life depression: a meta-analysis. *Int J Geriatr Psychiatr* 2003;18:1088-1094.
42. Hsieh C-J, Chang C, Su S-F, et al. Reminiscence Group Therapy on Depression and Apathy in Nursing Home Residents With Mild-to-moderate Dementia *J Exp Clin Med* 2010;2:72-78.

43. Leontjevas R, Gerritsen DL, Koopmans RTCM, et al. Process evaluation to explore internal and external validity of a multidisciplinary depression intervention in nursing homes. *J Am Med Dir Assoc* 2012;13(5):488.e1-8.
44. Starkstein SE, Jorge R, Mizrahi R, et al. A prospective longitudinal study of apathy in Alzheimer's disease. *J Neurol Neurosur Ps* 2006;77:8-11.

Chapter

9

Summary And General Discussion

This thesis addresses depression management in nursing homes. The primary aim was to evaluate the validity and effectiveness of the multidisciplinary intervention *Act in case* of Depression (AiD) in both somatic and dementia special care (DSC) units. AiD was developed in order to improve the detection and treatment of depression and depressive symptoms in nursing home (NH) residents. The program provides pathways for depression assessment procedures (detection, screening and diagnosing), and multidisciplinary treatment protocols (behavioral, psychosocial and pharmacological).

The Introductory Section presents information on nursing homes in the Netherlands, depression figures, recognition and treatment of depressive features in NH residents, and introduces the AiD program (Chapters 1 and 2). The Results Section includes the validity appraisal (Chapters 3 to 6), and the effect analyses (Chapters 7 and 8). This final chapter summarizes the main results, provides methodological considerations, and draws conclusions with implications for future research and daily practice.

The research questions in this thesis were, for validity appraisal:

1. Is the Nijmegen Observer-Rated Depression (NORD) scale a valid instrument for screening for depression in residents with and without dementia?

2. Is the Cornell Scale for Depression in Dementia (CSDD) an accurate proxy-based instrument in dementia when professional caregivers are the only available source of information?

3. Is the abbreviated 10-item Apathy Evaluation Scale a valid instrument to assess apathy in nursing home residents?

4. Is AiD feasible in nursing homes, and can data of the AiD trial be used for analyzing the effectiveness of the program?

The research questions for AiD effectiveness were:

5. Is the AID care program effective regarding depression prevalence?

6. Does the AiD program influence apathy in NH residents?

Summary of main findings

The Nijmegen Observer-Rated Depression scale can be used for screening for depression on a regular basis in NH residents with and without dementia. In a cross-sectional sub-study among NH residents with and without dementia, the accuracy of a new brief depression instrument, the Nijmegen Observer- Rated Depression (NORD) scale, was tested against standardized diagnostic criteria for depression (Chapter 3). Depression was diagnosed in accordance with the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition, Revised (DSM-IV-TR)[1] in residents without dementia, and the Provisional Diagnostic Criteria for Depression in Alzheimer Disease (PDCdAD)[2] in residents with dementia. Primary professional caregivers scored the NORD. Five of the six proposed NORD items were considered useful for distinguishing between residents with and without depression. The five-item scale showed accuracy figures that are within the range of the figures found for more extensive instruments[3-5] (area under the curve [AUC] determined using receiver operating characteristic [ROC] analyses in dementia: 0.84; 95% Confidence interval [CI], 0.75 to 0.90); in non-dementia: 0.84; 95% CI, 0.73 to 0.91). The scale showed a low negative likelihood ratio of 0.0 in non-dementia and of 0.2 in dementia. Considering this, and the relatively short time to administer the scale (2.4 minutes on average), it was concluded that the NORD can be used on a regular basis in residents both without dementia and with dementia.

The Cornell Scale for Depression in Dementia has acceptable accuracy when professional caregivers are the only source of information. Chapter 4 presents the results of a cross-sectional validation study testing the accuracy of the Cornell Scale for Depression in Dementia[6] (CSDD) in residents with dementia. The 'gold standard' was depression diagnosed using the PDCdAD.[2] CSDD accuracy, determined using ROC analyses, was compared with the accuracy of the Montgomery-Åsberg Depression Rating Scale[7] (MADRS). Both scales were scored during an interview with professional caregivers. Given the fact that proxies may have difficulty evaluating non-observable depressive symptoms (e.g. presence of suicidal or pessimistic thoughts), four methods for imputation of missing items were tested: imputation of the (1) lowest possible item score, i.e. 0 for CSDD items; (2) middle-most item score, i.e. 1 for CSDD items; (3) highest item score, i.e. 2 for CSDD items; and (4) an average score of the remaining items. There was no significant difference between the MADRS and CSDD with regard to accuracy. For the CSDD, imputation of the lowest possible score for missing items showed significantly better AUCs than the other methods. This method revealed an AUC of 0.79 (95% CI, 0.69 to 0.86), which was virtually the same as that in a previous study in which the score was based on information from

interviews of both residents and their proxies.[5] It was concluded that the CSDD has acceptable accuracy in NH residents with dementia when professional caregivers are the only available source of information.

The abbreviated Apathy Evaluation Scale is a valid instrument to assess apathy in NH residents. Chapter 5 presents the results of a validation study in which the Apathy Evaluation Scale[8] (AES-10) and the Neuropsychiatric Inventory apathy subscale[9] (NPIa) were compared on their discriminant validity, and on their performance to distinguish residents as apathetic or non-apathetic. Apathy was diagnosed using the Diagnostic Criteria for Apathy in Alzheimer's Disease and other Neuropsychiatric Disorders (DCA).[10] The AES-10 and NPIa correlated significantly and moderately with each other (r_s =0.62), which implies that their congruent validity was acceptable. The AES-10 correlated weakly (r_s, 0.27) and the NPIa moderately (r_s, 0.46) with the CSDD, which implies that the AES-10 may be preferred regarding discriminant validity. No significant difference was found between the instruments in their ability to distinguish apathetic from non-apathetic NH residents. The AES-10 produced higher sums of sensitivity and negative predictive value than the NPIa, which implies that the AES-10 can be preferred for screening purposes. Explorative accuracy analyses revealed that AUCs were significant in non-dementia (AES-10: AUC, 0.88; NPIa: AUC, 0.77), but not in dementia. It was concluded that both the AES-10 and NPIa may be used to distinguish apathetic from non-apathetic residents in a heterogeneous sample with and without dementia, or in residents without dementia.

A new process evaluation model is presented for providing insight into internal and external validity of the trial. In Chapter 6, AiD process data on sampling quality and intervention quality were evaluated for the purpose of gaining insight into internal and external validity of the intervention. According to our model, data on sampling quality (the procedures and results of the recruitment and randomization, as well as the maintenance of the participants) and intervention quality (relevance and feasibility, and the extent to which the complex intervention was performed as planned) constitute the first-order process data. In contrast to the second-order process data, which entail describing the delivered and received implementation strategies and the barriers and facilitators to carrying out the program, the first-order process data should be analysed prior to analyses of the effectiveness.

Sampling quality evaluation showed the need to account for the type of unit in effect analyses. Data showed that, although including the units by convenience could influence external validity, other recruitment strategies, such as no exclusion criteria for the residents, and ongoing recruitment of new residents, were used to improve both

external and internal validity. There was a different pattern for the informed consent (IC) rate between DSC and somatic units, which implies the need to account for the type of unit in effect analyses. While the IC rate in DSC units declined from 62% at baseline to 50% at the end of the study, despite of recruitment of newly admitted residents, the rate in somatic units improved gradually from 52% to 56% at the fifth measurement and dropped to 49% at the last (sixth) measurement. The IC rates were comparable with the rates reported in other studies on depression management.[11,12]

Intervention quality evaluation showed that AiD is feasible. NH staff members involved in carrying out the program were satisfied with AiD and reported that the program was feasible and relevant (external validity).

Intervention quality evaluation showed the need to account for the extent to which AiD elements were performed in effect analyses. DSC and somatic units differed in the extent to which several AiD components were performed. Furthermore, the average proportion of residents that received AiD treatment components when indicated was 40%. This was lower than 76% for AiD assessment components. The first screening phase was performed by nursing staff in all units. Monitoring of the treatment results was performed least often (18%). Because this might influence internal validity (the effects could be attributable to AiD components performed and not to the whole program), it was concluded that effect analyses should not only account for the type of unit but also for the extent to which AiD components were performed.

AiD improved depression outcomes in somatic residents, and the quality of life in DSC and somatic residents. The primary effect analyses showed that the AiD program reduced depression prevalence in somatic units (endpoint effect estimate, percent points, -7.3; 95% CI, -13.7 to -0.9; Cohen's d, -0.6) (Chapter 7). The results of the secondary analyses showed that the severity of depressive symptoms decreased in somatic units (CSDD points, -0.8; 95% CI, -1.4 to -0.1), and quality of life (QoL) as reported by the residents improved due to the program in both DSC and somatic units (visual analogue scale of the Euroqol-5D[13] points, 3.4; 95% CI, 0.5 to 6.3). Additional analyses on adherence rates for AiD algorithms revealed that effect-sizes for outcomes on which the intervention as a whole had a significant effect were larger when assessment procedures were applied (endpoint Cohen's d, -1.1) but not when treatment algorithms were applied, which were implemented poorly. In residents without severe cognitive impairments, no intervention effect was found using the 8-item Geriatric Depression Scale[14] (GDS8). This could be due to power loss, because when residents with severe cognitive impairments were included in analyses, there was an indication that GDS8 scores decreased in somatic

units in the intervention condition. Proxies evaluated QoL of the residents lower in the intervention condition, which can be due to raised caregivers' awareness concerning the influence of depression on QoL. When AiD treatment was provided, caregivers evaluated QoL of the residents as improved. It was concluded that the program can be used for reducing depression prevalence and the severity of depressive symptoms, and improving QoL in somatic units. The program improved QoL in DSC units, which indicated that AiD can also be used effectively in DSC units. Better implementation of treatment protocols in daily practice may improve AiD effects on depression and quality of life.

For reducing apathy and depressive motivational symptoms, activating strategies can be used effectively in DSC units, and psychotherapy in somatic residents. Chapter 8 presents the results of the AiD effect analyses on apathy outcomes (AES-10) and two groups of depressive symptoms, i.e. motivational (CSDD items retardation, loss of interest, appetite loss, weight loss, and lack of energy), and mood symptoms (CSDD items anxiety, sadness, irritability, agitation, diurnal variation, suicide, poor self-esteem, and pessimism). The whole AiD program reduced apathy (AES-10 points, -2.3; 95% CI, -3.3 to -1.3) in DSC units and depressive motivational symptoms in somatic units (CSDD motivational subscale points, -0.4; 95% CI, -0.7 to - 0.1). Mood symptoms were not affected by AiD. Exploring the effects of the individual AiD components, we revealed that in DSC units, behavioral activation strategies (day program and pleasant activities plan, treatment module 1) were the major contributors to the positive effect on apathy (-4.8; 95% CI, -6.7 to -2.9; Cohen's d, -0.73), and also were associated with reduced motivational symptoms (-0.9; 95% CI, -1.3 to - 0.4; Cohen's d, -0.45). In somatic units, psychotherapy (treatment module 2) was the major contributor to the effect on motivational symptoms (-0.8; 95% CI, -1.4 to - 0.3; Cohen's d, -0.80), and also was associated with reduced apathy (-3.2; 95% CI, -5.1 to -1.3; Cohen's d, -0.42). Pharmacological treatment of depression was associated with worsening apathy in both DSC and somatic units (1.8; 95% CI, 0.5 to 3.2) but did not show effects on motivational or mood symptoms. It was concluded that the AiD program can be used for reducing apathy, and motivational depressive symptoms. Activating strategies can be especially effective in DSC residents, and psychotherapy in somatic residents.

Methodological considerations

Design

The AiD trial is a pragmatic clinical trial. In contrast to an explanatory trial, which is conducted to determine efficacy under standardized conditions, a pragmatic clinical trial

measures effectiveness, i.e. the effects in 'real-life' conditions. Because interventions in a pragmatic trial are applied flexibly as they would be in normal practice,[15] the results are more easily translated into practice than the results of an explanatory trial.

For answering the research questions, a pragmatic longitudinal trial with a stepped-wedge cluster-randomized design[‡] was performed resulting in two longitudinal studies on AiD effectiveness, and three cross-sectional validity sub-studies using baseline data. Alongside the AiD trial, process data were collected and evaluated in order to understand internal and external validity of the intervention.

A stepped-wedge design was used because of the pragmatic nature of the trial, and practical considerations. There can be several reasons to perform a stepped-wedge RCT instead of a parallel RCT. First, the stepped-wedge design is particularly useful for pragmatic trials, because the clusters' gradual crossing-over from the control to the intervention condition allows fine-tuning of the implementation strategies delivery.[16] In order to facilitate program delivery by stakeholders, various implementation strategies were used (see Information Box 9.1) and these were fine-tuned (e.g. used forms) after the first AiD intervention group crossed over to the intervention condition. On the whole, the design ensured the provision of the same implementation strategies to all intervention groups during the study.

Second, providing the intervention in all participating units was expected to increase the motivation of NH staff and management for participating in the study.[17] The stepped-wedge design ensured that all residents would receive depression care according to the intervention before the end of the study. This is an important ethical consideration for interventions that are believed to be superior compared to the control condition.[16]

Third, this design maximizes statistical power, and may require fewer clusters compared with a parallel cluster RCT.[16] From the practical perspective, fewer clusters and gradual crossing-over to the intervention condition by the intervention groups improved research management and should reduce trial expenditures.

Cluster randomization was performed to avoid contamination bias. To avoid contamination between subjects within one unit a cluster-randomized design was used in the trial. However, in choosing the cluster randomization we might increase confounding because we could not account for the allocation of individual characteristics over the

[‡] The stepped-wedge design is a variant of a one-way crossover cluster trial design in which all clusters receive the intervention but the moment when this happens is randomly ordered. See Chapter 2 and Chapter 6 for more information.

intervention groups. Yet, all clusters in the stepped-wedge design crossed over from the control to the intervention condition, which should minimize confounding. Most residents' characteristics would remain the same in the control and intervention conditions, whereas including new residents during the study would minimize time dependent characteristics, such as the severity of an illness.

Implementation strategies

It is not clear whether the used implementation strategies may have influenced the results independently of AiD components. In the effectiveness trial, implementation strategies were used (Information box 9.1) for introducing the care program in the units. These strategies, however, may have influenced the outcomes in the trial independently of AiD components. For example, we found an effect of the assessment procedures in general (Chapter 7), and individual assessment phases in particular (Chapter 8) on different outcomes. This may reflect effects of the implementation strategies used, and the training of the staff could play a specific role. By providing the same strategies to all intervention groups, which was considered necessary to improve internal validity, we limited the possibility to test effects of individual implementation strategies. Second-order process data on the implementation strategies may provide qualitative and quantitative input for further analyses. The analysis of these data is beyond the scope of this thesis.

Information box 9.1

Implementation strategies used

For units crossing-over from the control to the intervention condition, the research team used three strategies. First, for providing information, the research team delivered printed and digital material tailored for the stakeholders: psychologists, physicians, and nursing staff, including the unit manager. Second, education was provided to the nursing staff (a 3.5 hours course) about depressive symptoms and the AiD program. The life review therapy training[18] was provided to the psychologists of the units that were about to cross-over to the intervention condition. A medication protocol was provided to the unit physician, who was contacted by phone to discuss the protocol with a physician involved in the protocol development. Third, tailored communication was provided including semi-structured 'content' and 'feedback' interviews. In a 'content interview', physicians and psychologists were invited to discuss possible questions about the program content. For a 'feedback interview', the GDS8 and CSDD scores, administered by the research team during regular measurements, were provided to psychologists at least once during the intervention condition. The scores were discussed as well as treatment modules used by the multidisciplinary teams. The psychologists could use the data on the GDS8 and CSDD for phase 2 (screening) if they still had not performed the screening phase. This could facilitate providing other AiD components including treatment of residents with depression or depressive symptoms without severe delays. The three strategies were used in all intervention groups to foster the use of AiD components.

Changes in AiD protocol

A protocol change was made for the use of depression scales. The initial protocol prescribed the use of the Brief Observer-Rated Screening Scale for Depression in Elderly Medical Patients (Hammond-scale).[19] However, this scale was not developed or tested for NH residents with and without dementia. We developed the NORD scale based on the Hammond-scale, and the results of the validity study (Chapter 3) showed that the scale can be used in residents with and without dementia. Therefore, the NORD was used in the first screening phase of AiD.

According to the initial trial protocol (Chapter 2), the GDS8 should be used in somatic units and the CSDD in DSC units for depression outcomes. The protocol was changed after the first week of the measurements: both scales were further used in both unit types. We encountered GDS8 scoring problems due to language or cognitive problems more often than expected in somatic residents. In somatic residents with language problems, the CSDD could be administered and used for imputation of the GDS8 scores. However, the

interviewers reported a poor match between the GDS8 and CSDD, which was confirmed later by the scales' weak correlation (Chapter 7). A second reason to change the protocol was that it was considered important to use the same scales in both DSC and somatic residents for the comparison of the effects between the two unit types.

Depression

Depression scales were used instead of standardized diagnostic criteria. Unfortunately, using a comprehensive diagnostic procedure was not possible in our study without breaking the NH professionals' blinding for the outcomes. Therefore, depression prevalence was assessed using the cut-off scores of depression scales. By applying the CSDD cut-off scores that are widely accepted in depression literature, we enabled other researchers to use our results for comparison with other (future) studies. For the GDS8, we choose cut-off scores based on a validation study using the same scoring procedure as ours.[20] Depression prevalence based on the CSDD (46%) and GDS8 (32%) scales in the AiD trial may be considered within the range of 20 to 68% reported in various NH studies using GDS versions or the CSDD.[21-28]

The same scoring method for the CSDD was used in DSC and somatic units. In this thesis, the CSDD was additionally validated for scoring the scale when only one source of information is available in residents with dementia. We used the same scoring method in somatic residents. Although using different sources of information when assessing depression should be preferred in clinical practice, particularly in residents without severe cognitive impairments, using the same instrument and the same scoring method was considered important in the trial. Unfortunately, due to the initial protocol, not enough CSDD data were collected prior to the intervention for additional validation of the CSDD in somatic residents. Yet, considering that cognitive problems are less prevalent in somatic residents, that the CSDD was validated in persons without severe cognitive problems,[29] and that reports of proxies tend to be consistent with reports of elderly people without cognitive impairments when the same depression instrument is used,[30] we believe that the CSDD is also valid in somatic residents when only one source of information is used, namely professional caregivers who are in daily contact with the resident.

Lack of AiD effect on the GDS8 outcomes was probably due to the scale's emphasis on mood. It was hypothesized in the Discussion section of Chapter 7 that one of the reasons why an AiD effect on GDS8 outcomes was not found was because the GDS8 represents mood symptoms only. In contrast, the CSDD includes not only mood symptoms, but also other aspects such as motivational symptoms. The second effect study (Chapter

8) confirmed this assumption as the AiD program and its individual components did not influence CSDD mood symptoms, but did influence motivational symptoms.

Apathy

Revealed AiD effects underpin construct validity of apathy measured by the AES-10. The effect analyses showed that AES-10 scores decreased due to the whole AiD program in DSC residents (Chapter 8). This decrease was mainly attributable to activating strategies. This may support the construct validity of the AES-10 because activation strategies have been associated with reducing apathy.[31] It is not surprising that the effect-size of psychotherapy was larger on depressive motivational symptoms than on apathy in somatic residents: the primary goal of psychotherapy was to reduce depressive symptoms. Unfortunately, apathy worsened in those units where a medication protocol for treatment of depression was used. This was the case in both DSC and somatic units, which, again, underpins the validity of AES-10 because antidepressant treatment may induce apathy.[32,33]

Implications for future research

Using a stepped-wedge design

It appeared that some participating residents and NH staff members experienced the six measurements of the trial as burdening. **Using a stepped-wedge design with fewer measurements should be considered in future research.** Furthermore, it is possible that the four-month period between measurements was not sufficient to reach effect on the outcomes. Yet, only the last of five intervention groups had just one follow-up after the start of the intervention period, so the effect of this was likely to be small in the AiD trial. Nevertheless**, a design with a period longer than four months between measurements may be preferred for a complex health care program.**

First-order process evaluation

In contrast to an explanatory trial, conditions in a pragmatic trial are not ideal or strictly selected, and that may influence intervention validity. Process evaluation can help gaining insight into internal and external validity of a complex intervention and that can help determining whether further statistical analyses are meaningful. Therefore, **for pragmatic trials, evaluation of first-order process data is advised prior to executing statistical analyses.**

Assessment

When comparing the accuracy of the tests across the full range of cut-offs using AUCs, the NORD (Phase 1 screening instrument) tended to have similar or even slightly better results than the CSDD (Phase 2 screening instrument)(Chapters 3 and 4). The CSDD did not show a high sum of specificity (recognizing residents without depression) and positive predictive value (minimizing false positive decisions), which is preferable for an instrument to be used in a second screening step (see Appendix A1.2). Therefore, the additional value of using the CSDD *after* using the NORD is questionable. In general, **research on the benefits of a two-step screening above a one-step screening procedure is needed.** To the best of our knowledge, research remains to be done that accounts for the benefits and costs of using different cut-off scores from different perspectives: that of the residents and of the health providers.[34]

The Apathy syndrome is not included in the glossary of traditional nomenclatures such as the DSM-IV-TR, and there is no clear consensus on the definition of apathy. The abbreviated AES-10 was specifically designed for NH residents. Possibly, the non-significant result for the AES-10 accuracy in residents with dementia can be explained by low numbers of residents with dementia in our validity study, or by poor validity of the newly proposed DCA criteria in NH residents. **Further validation of the AES-10, and validation of standardized criteria of apathy in residents with dementia is needed.**

There is no single definition for instrument validity. In this thesis, validity testing was limited to (1) congruent validity against scales measuring the same construct; (2) discriminant validity against scales measuring a different construct; and (3) concurrent validity (a form of criterion validity) against a 'gold standard'. **Appraisal of other types of validity is recommended.** Predictive validity, for example, refers to the extent to which a scale predicts the same construct in the future. Sensitivity to change of the CSDD and AES-10 were demonstrated in the effect analyses, but the extent to which the NORD is sensitive to change has not yet been studied. It is also unclear, whether, for the scoring methods used, the scales have an acceptable intra-individual and intra-observer variability. Therefore, **further research on reliability of the tests is recommended.**

Treatment

Considering that activating strategies (treatment module 1) were initiated less often in somatic residents than in DSC residents, **research on how activating strategies can be used effectively in somatic NH residents is advised.** The poor use of module 2 in DSC residents can reflect existing challenges associated with talk therapies in dementia.

However, psychologists also had a protocol for mediative behavioral therapy which they could use in case the advised Life Review therapy or other forms of psychotherapy could not be performed. **Research is needed to explore how NH psychologists can use ´talk therapies´ and mediative behavioral strategies more effectively in DSC units.** Next to treatment module 2 in DSC units, phase 5 monitoring was implemented least often in both DSC and somatic units. The UKON task force responsible for the AiD program considered this AiD component essential for depression management. In general, **more research is needed on how monitoring of treatment effects using standardized screening instruments in NHs can be improved.**

Generalizability

Given the effects of AiD on depression, quality of life, and apathy outcomes, **further evaluation is recommended into barriers and facilitators** to carrying out the AiD program. Once cost-effectiveness has been determined, and AiD improvements have been made, it is further recommended to explore ways of implementing the program successfully in nursing homes. Because the trial was performed in UKON nursing homes, and the program and instruments were in Dutch, **the generalizability of the results should be determined in nursing homes other than UKON and in other countries.** In addition, **replication of the studies is needed in long term care facilities other than nursing homes.**

Implications for NH practice

Screening in practice

The validity studies for the NORD, CSDD and AES-10 provided cut-off scores that can be used by NH professionals. Choosing a cut-off score should depend on the expected advantages and disadvantages of correct and incorrect decisions. Information box 10.1 (Appendix 1) presents several cut-off scores that can be used in practice. The results demonstrated that **the scales can be used for ruling out a condition.** Of the scales evaluated, the NORD showed promising results for ruling in the condition (confirming the diagnosis). However, this was found for residents in dementia only and the results should be replicated before any recommendations can be made. In sum, **using the scales without further comprehensive diagnostic procedure is not recommended,** especially for starting treatment.

Multidisciplinary teams

The multidisciplinary care of Dutch nursing homes provides excellent prerequisites for managing depression effectively. **Involving nursing staff in recognizing depression is the**

first step for improvement of depression care in NHs. By performing the first screening step on a regular basis in the AiD trial, nursing staff demonstrated that they can play an important role in depression care. **The unit psychologist and physician should initiate and carry out further depression assessment and treatment** after the nursing staff provides information about depressive symptoms in residents.

Although the coordinating role of the psychologist is described explicitly in the program, psychologists of one-third of the participating units did not play an active role in coordinating and carrying out the program, neither did they support nursing staff in introducing activating strategies (data to be published). **Psychologists should support and stimulate nursing staff in performing activating strategies.** In order to improve effectiveness of activating strategies (AiD treatment module 1), particularly in somatic residents, **NH staff should be able to tailor the strategies to the needs and abilities of the resident.** A crucial factor for successful application of module 1 is the **involvement of the resident's relatives** in facilitating and executing the pleasant-activities plan.[35] Therefore, implementation strategies such as tailored information about activating strategies should also be provided to the relatives. Another important factor is **finding a local opinion leader,**[35] who may help the psychologist in coordinating the program.

Given the limited evidence for the effects of antidepressants in NHs, and no effect found in this trial for the use of a medication protocol, **physicians should not rely on pharmacological treatment alone** without first considering psychosocial treatment of depression. The revealed effects of psychosocial AiD components underpin that **a multidisciplinary approach is needed.**

Nursing home management

Frequent staff changes in the units not only led to additional research-team burden and expenditures, but also compromised performing the program as planned. Out of 33 units in total, 22 physicians, 19 psychologists, and 11 unit managers accepted another job or left the unit during the intervention condition (during an average 12 month period). In 12 units, more than 20% of the nursing staff educated by the researchers left the unit. It is not surprising, then, that the NH professionals mentioned frequent staff changes and shortage of (qualified) staff as an important barrier to carrying out the program (preliminary exploration of the second-order process data, not reported). **Supporting stable multidisciplinary staff and empowering personnel are crucial factors for introducing care improvement programs.** According to the NH professionals, implementation would have been facilitated if - in addition to less staff turnover - more would be invested in educational strategies and in activities for residents. **Permanent**

investment in the education of NH staff about depression and in activities for the residents should always be high on the agenda of NH management.

Policy makers

Given that depression care is a quality indicator according to the national Quality Framework for Responsible Care,[36] **policy makers should create favorable conditions for NH professionals to improve NH depression care.** Unit managers of different units reported cuts in nursing staff and recreational therapists, and cancelation of recreational activities due to budget restrictions. Obviously, it is difficult to apply psychosocial components in these circumstances. **Policy makers need to ensure that** – even in times of budget cuts – **effective psychosocial strategies can be used in NHs,** because pharmacological treatment alone is less effective than a multidisciplinary approach including psychosocial interventions.

To reduce staff turnover and to ensure stable and qualified personnel, **the NH sector should be made more attractive for qualified professionals.** Innovative evidence-based care approaches have been developed and working according to the 'state of the art' approaches appeals to many professionals. However, **to deliver high quality care, professionals must be adequately equipped.** For instance, knowing which strategies are effective for depression recognition and treatment is important, but being unable to execute such strategies because of budget restrictions, high rates of reorganizations, high staff turnover or a shortage of staff, is frustrating for NH professionals. These adverse circumstances will not only result in poorly treated depression, but will also discourage professionals from starting or continuing a career in the NH sector.

Concluding remarks

To date, only few complex intervention studies have been attempted in a NH setting - a setting that shows high rates of reorganization and staff turnover; low nursing staff education level, and a shortage of staff. Taking on this challenge and facing encountered problems was a major strength of this pragmatic trial. In addition to the useful information on the instruments tested, **the AiD trial provided evidence for the effectiveness of the depression care program** in Dutch nursing homes. Unfortunately, the AiD trial also **showed that depression care in Dutch nursing homes needs to be further improved.** A proportion of the residents was neither screened nor treated using behavioral and psychosocial strategies, even in the intervention condition. The current economic challenges may delay the improvement of depression care because financial cuts in nursing care may result in limited staff time. Nevertheless, it is not always a matter of

significant time investment by staff when improvement of (depression) care is needed. Some of the AiD components, such as the first screening step using the NORD, require little staff time, and the routine use of other components will reduce time investment because of the growing staff expertise. It is not about what we cannot do to reduce emotional suffering of clients. It is about what we can do:

AiD: we can Act in case of Depression!

Literature

1. APA. Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association, 2000.
2. Olin JT, Schneider LS, Katz IR, et al. Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002;10:125-128.
3. Jongenelis K, Pot AM, Eisses AM, et al. Diagnostic accuracy of the original 30-item and shortened versions of the Geriatric Depression Scale in nursing home patients. *Int J Geriatr Psychiatry* 2005;20:1067-1074.
4. Leontjevas R, van Hooren S, Mulders A. The Montgomery-Asberg Depression Rating Scale and the Cornell Scale for Depression in Dementia: a validation study with patients exhibiting early-onset dementia. *Am J Geriatr Psychiatry* 2009;17:56-64.
5. Barca ML, Engedal K, Selbaek G. A reliability and validity study of the cornell scale among elderly inpatients, using various clinical criteria. *Dement Geriatr Cogn Disord* 2010;29:438-447.
6. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23:271-284.
7. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Brit J Psychiat* 1979;134:382-389.
8. Lueken U, Seidl U, Volker L, et al. Development of a short version of the Apathy Evaluation Scale specifically adapted for demented nursing home residents. *Am J Geriatr Psychiatry* 2007;15:376-385.
9. Cummings JL, Mega M, Gray K, et al. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. *Neurology* 1994;44:2308-2314.
10. Robert P, Onyike CU, Leentjens AF, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. *Eur Psychiatry* 2009;24:98-104.
11. Brodaty H, Draper BM, Millar J, et al. Randomized controlled trial of different models of care for nursing home residents with dementia complicated by depression or psychosis. *J Clin Psychiat* 2003;64:63-72.
12. Rosen J, Rogers JC, Marin RS, et al. Control-relevant intervention in the treatment of minor and major depression in a long-term care facility. *Am J Geriat Psychiat* 1997;5:247-257.
13. Group E. EuroQol - a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199-208.
14. Jongenelis K, Gerritsen DL, Pot AM, et al. Construction and validation of a patient- and user-friendly nursing home version of the Geriatric Depression Scale. *Int J Geriatr Psychiatry* 2007;22:837-842.

15. Zwarenstein M, Treweek S, Gagnier JJ, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ* 2008;337:a2390.
16. Mdege ND, Man MS, Taylor Nee Brown CA, et al. Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* 2011;64:936-948.
17. Gerritsen DL, Smalbrugge M, Teerenstra S, et al. Act In case of Depression: The evaluation of a care program to improve the detection and treatment of depression in nursing homes. Study Protocol. *BMC Psychiatry* 2011;11:91.
18. Bohlmeijer E, Steunenberg B, Leontjevas R, et al. In Ducth: [Dierbare Herinneringen: Protocol voor Individuele life-review therapie gebaseerd op autobiografische oefening]. Enschede: Universiteit Twente, 2010.
19. Hammond MF, O'Keeffe ST, Barer DH. Development and validation of a brief observer-rated screening scale for depression in elderly medical patients. *Age Ageing* 2000;29:511-515.
20. Smalbrugge M, Jongenelis L, Pot AM, et al. Screening for depression and assessing change in severity of depression. Is the Geriatric Depression Scale (30-, 15- and 8-item versions) useful for both purposes in nursing home patients? *Aging Ment Health* 2008;12:244-248.
21. Snowdon J, Fleming R. Recognising depression in residential facilities: An Australian challenge. *Int J Geriatr Psychiatry* 2008;23:295-300.
22. Barca ML, Selbaek G, Laks J, et al. Factors associated with depression in Norwegian nursing homes. *Int J Geriatr Psychiatry* 2009;24:417-425.
23. McCurren C, Dowe D, Rattle D, et al. Depression among nursing home elders: testing an intervention strategy. *Applied nursing research* 1999;12:185-195.
24. Damian J, Pastor-Barriuso R, Valderrama-Gama E. Descriptive epidemiology of undetected depression in institutionalized older people. *J Am Med Dir Assoc* 2010;11:312-319.
25. Katz IR, Lesher E, Kleban M, et al. Clinical features of depression in the nursing home. *Int Psychogeriatr* 1989;1:5-15.
26. Drageset J, Eide GE, Ranhoff AH. Depression is associated with poor functioning in activities of daily living among nursing home residents without cognitive impairment. *J Clin Nursing* 2011;20:3111-3118.
27. Harwood DG, Ownby RL, Barker WW, et al. The factor structure of the Cornell Scale for Depression in Dementia among probable Alzheimer's disease patients. *Am J Geriatr Psychiatry* 1998;6:212-220.
28. Smalbrugge M, Jongenelis L, Pot AM, et al. Incidence and outcome of depressive symptoms in nursing home patients in the Netherlands. *Am J Geriatr Psychiatry* 2006;14:1069-1076.
29. Alexopoulos GS, Abrams RC, Young RC, et al. Use of the Cornell scale in nondemented patients. *J Am Geriatr Soc* 1988;36:230-236.
30. Chen CY, Liu CY, Liang HY. Comparison of patient and caregiver assessments of depressive symptoms in elderly patients with depression. *Psychiatr Research* 2009;166:69-75.
31. Brodaty H, Burns K. Nonpharmacological Management of Apathy in Dementia: A Systematic Review. *Am J Geriatr Psychiatry* 2012;20(7):549-564.
32. Barnhart WJ, Makela EH, Latocha MJ. SSRI-induced apathy syndrome: a clinical review. *J Psychiatr Pract* 2004;10:196-199.
33. Settle EC, Jr. Antidepressant drugs: disturbing and potentially dangerous adverse effects. *J Clin Psychiatr* 1998;59 Suppl 16:25-30; discussion 40-22.
34. Smits N, Smit F, Cuijpers P, et al. Using decision theory to derive optimal cut-off scores of screening instruments: an illustration explicating costs and benefits of mental health screening. *Int J Method Psych* 2007;16:219-229.
35. Verkaik R. Depression in dementia. Development and testing of a nursing guideline. [PhD thesis]. Utrecht: Universiteit Utrecht, 2009.

36. ActiZ A, LOC, et al. Quality Framework Responsible Care. Rijswijk: ActiZ, 2007.
37. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003;29:1043-1051.

Chapter

10

Nederlandse Samenvatting [Dutch Summary]

Vraagstellingen

De focus van dit proefschrift ligt op de verbetering van de depressiezorg in verpleeghuizen. Het aandeel mensen met een depressie is veel hoger in verpleeghuizen dan in de thuissituatie. Depressie heeft negatieve consequenties voor het welbevinden van verpleeghuiscliënten en wordt zowel met een hoger zorggebruik als met een hogere mortaliteit geassocieerd. Desondanks wordt deze aandoening in verpleeghuizen onvoldoende opgespoord en behandeld. Het structureel opsporen van depressieve klachten met gevalideerde screeningsinstrumenten komt niet of nauwelijks voor. Terwijl uit onderzoek blijkt dat een depressiebehandeling het meest effectief is wanneer naast medicatie ook psychosociale interventies worden toegepast, blijft een depressie-behandeling in verpleeghuizen nog vooral medicamenteus.

De aanwezigheid van verschillende disciplines in Nederlandse verpleeghuizen biedt uitstekende mogelijkheden om depressies op een effectieve manier op te sporen en te behandelen. Met het doel de depressiezorg in verpleeghuizen te verbeteren heeft het Universitair Kennisnetwerk Ouderenzorg Nijmegen (UKON[§]) een multidisciplinair zorgprogramma ontwikkeld: *Doen bij Depressie* ('Act in case of Depression'). Aangezien depressie onvoldoende wordt herkend, is het van groot belang dat multidisciplinaire teams niet alleen over procedures beschikken om depressies te behandelen, maar ook over structurele procedures om depressieve klachten op te sporen. *Doen bij Depressie* biedt dan ook een direct implementeerbaar werkplan om depressieve klachten en depressies zowel op te sporen als te behandelen. Het is gebaseerd op nationale en internationale richtlijnen voor depressiezorg en op evidentie uit wetenschappelijk onderzoek. In de periode van mei 2009 tot en met april 2011 is een multi-center gerandomiseerd gecontroleerd onderzoek uitgevoerd om de effectiviteit van het werken volgens *Doen bij Depressie* in de dagelijkse praktijk te testen.

Dit proefschrift bestaat uit drie delen. Het inleidende deel omvat twee hoofdstukken. Hoofdstuk 1 presenteert inleidende informatie over verpleeghuizen in Nederland, depressiecijfers en het opsporen en behandelen van depressieve klachten bij verpleeghuiscliënten. Er wordt ingegaan op het belang van een multidisciplinair zorgprogramma en op de kenmerken van *Doen bij Depressie*. In hoofdstuk 2 wordt het onderzoeksprotocol gepresenteerd. In het tweede deel van dit proefschrift komen de

[§] UKON is een samenwerkingsverband tussen de afdeling Eerstelijnsgeneeskunde van het Universitair Medisch Centrum St Radboud en meerdere organisaties voor ouderenzorg en chronisch zieken (voor recente informatie over deelnemende organisaties zie www.uko-n.nl).

resultaten van deelstudies aan bod. Voorafgaand aan de effectiviteitanalyses zijn drie deelstudies uitgevoerd naar de validiteit van meetinstrumenten die gebruikt worden in het zorgprogramma en in de effectiviteitanalyses (hoofdstukken 3, 4 en 5). Verder worden de resultaten van een procesevaluatie gepresenteerd die uitgevoerd is om te bepalen of effectiviteitanalyses zinvol zijn (hoofdstuk 6). De resultaten van twee effectiviteitstudies staan in de hoofdstukken 7 en 8 beschreven. In het laatste deel wordt een samenvatting gegeven van de belangrijkste resultaten, worden methodologische overwegingen besproken en worden conclusies getrokken en implicaties voor toekomstig onderzoek en de dagelijkse praktijk beschreven (hoofdstuk 9).

De onderzoeksvragen voor de beoordeling van de validiteit zijn:

1. Is de Nijmegen Observer-Rated Depression (NORD) schaal een valide meetinstrument voor de screening op depressies bij cliënten met en zonder dementie? Voor het screenen van depressies bij verpleeghuiscliënten op een regelmatige basis is een snel in te vullen meetinstrument nodig dat bij mensen met en zonder dementie is toe te passen. De NORD is in het *Doen bij Depressie* onderzoek ontwikkeld om toe te passen als de eerste stap van *Doen bij Depressie*.

2. Is de Cornell Scale for Depression in Dementia (CSDD) een nauwkeurig meetinstrument bij dementie wanneer professionele zorgverleners de enige beschikbare bron van informatie zijn? Veel verpleeghuiscliënten met cognitieve problemen kunnen niet geïnterviewd worden terwijl de oorspronkelijke CSDD gescoord wordt op basis van informatie uit meerdere bronnen inclusief de cliënt zelf. De schaal werd gebruikt in de eerste effectiviteitstudie van dit proefschrift als centrale uitkomstmaat en fungeert als tweede stap van *Doen bij Depressie*.

3. Is de verkorte Apathie Evaluatie Schaal met 10 items (AES-10) een valide meetinstrument om apathie bij verpleeghuiscliënten te beoordelen? De AES-10 werd in dit proefschrift gebruikt om te bepalen of het zorgprogramma van invloed is op apathie.

4. Is het *Doen bij Depressie* zorgprogramma toepasbaar in verpleeghuizen, en kunnen gegevens van het *Doen bij Depressie* onderzoek worden gebruikt voor het analyseren van de effectiviteit van het zorgprogramma? Het is belangrijk om voorafgaand aan de effectiviteitanalyses naar de interne en externe validiteit van de interventie te kijken. Hiervoor werden procesgegevens geanalyseerd.

De onderzoeksvragen voor de effectiviteit van het *Doen bij Depressie* zorgprogramma waren:

5. Verlaagt het werken volgens het *Doen bij Depressie* zorgprogramma de prevalentie van depressie in verpleeghuiscliënten? De hoofdvraag van het *Doen bij Depressie* onderzoek is of het werken volgens het zorgprogramma de prevalentie van depressie verlaagt. Secundaire uitkomstmaten (de ernst van depressieve symptomen en de kwaliteit van leven) werden gebruikt ter ondersteuning van de hoofdanalyses.

6. Heeft het werken volgens *Doen bij Depressie* invloed op apathie bij verpleeghuiscliënten? Het zorgprogramma bevat behandelstrategieën die ook van invloed kunnen zijn op apathie.

Samenvatting van de bevindingen

De Nijmegen Observer-Rated Depression schaal kan gebruikt worden voor gestructureerde screening op depressie bij verpleeghuiscliënten met en zonder dementie. Hoofdstuk 3 introduceert een nieuwe korte schaal, de Nijmegen Observer-Rated Depression scale (NORD). De schaal werd getoetst ten opzichte van gestandaardiseerde diagnostische criteria voor depressie. Bij cliënten zonder dementie werd depressie gediagnosticeerd aan de hand van de *Diagnostic and Statistical Manual*, vierde editie, herzien (DSM-IV-TR). Bij cliënten met dementie gebeurde dit aan de hand van de *Provisional Diagnostic Criteria for Depression in Alzheimer Disease* (PDCdAD). Eerst verantwoordelijk verzorgenden vulden de NORD schaal in. De nauwkeurigheid van de NORD met vijf items bleek vergelijkbaar te zijn met langere screeningslijsten zoals de Geriatric Depression Scale en de Cornell Scale for Depression in Dementia. Voor de nauwkeurigheid werd de oppervlakte onder de curve (AUC) bepaald aan de hand van de Receiver Operating Characteristic (ROC) analyse: de AUC bij cliënten met dementie was 0,84 (95% betrouwbaarheidsinterval [BI]: 0,75 tot 0,90) en bij cliënten zonder dementie 0,84 (95% BI: 0,73 tot 0,91). De schaal bleek een lage negatieve likelihood ratio te hebben van 0,0 bij cliënten zonder dementie en van 0,2 bij cliënten met dementie. Gezien deze resultaten en de relatief korte tijd die het kost om de schaal in te vullen (gemiddeld 2 à 3 minuten), werd geconcludeerd dat de NORD gebruikt kan worden bij een regelmatige screening van zowel cliënten met als cliënten zonder dementie.

De Cornell Scale for Depression in Dementia heeft een acceptabele nauwkeurigheid wanneer professionele zorgverleners de enige bron van informatie zijn. Hoofdstuk 4 beschrijft een cross-sectionele validatiestudie van de Cornell Scale for Depression in Dementia (CSDD) bij cliënten met dementie. De 'gouden

standaard' was een depressie gediagnosticeerd aan de hand van de PDCdAD. De nauwkeurigheid van de CSDD, die met ROC-analyses werd bepaald, werd vergeleken met de nauwkeurigheid van de Montgomery-Åsberg Depression Rating Scale (MADRS). Beide schalen werden gescoord tijdens een interview met professionele zorgverleners. Aangezien de medewerkers problemen kunnen hebben bij het evalueren van niet-observeerbare depressieve symptomen zoals de aanwezigheid van suïcidale of pessimistische gedachten, werden vier methoden voor het opvullen (imputatie) van ontbrekende items getest: imputatie van (1) de laagst mogelijke itemscore, (2) de gemiddelde itemscore, (3) de hoogst mogelijke itemscore, en (4) de gemiddelde score van de overige items. Er is geen significant verschil gevonden tussen de nauwkeurigheid van de MADRS en de CSDD. Het imputeren van de laagst mogelijke itemscore (=0) gaf voor de CSDD een significant hogere AUC dan andere imputatiemethoden. Deze methode gaf een AUC van 0,79 (95% BI: 0,69 tot 0,86), wat vrijwel gelijk is aan een eerder gevonden AUC in een recente studie waarin zowel de cliënten als hun zorgverleners werden geïnterviewd. De conclusie is dat de CSDD gebruikt kan worden voor verpleeghuiscliënten met dementie wanneer slechts één bron van informatie beschikbaar is, namelijk die van professionele zorgverleners.

De verkorte Apathy Evaluation Scale is een valide meetinstrument om apathie te beoordelen bij verpleeghuiscliënten. Hoofdstuk 5 beschrijft een validatiestudie waarin de verkorte Apathy Evaluation Scale (AES-10) en de Neuropsychiatric Inventory apathy subscale (NPIa) werden vergeleken op hun discriminante validiteit, en op hun vermogen om cliënten met of zonder apathie te onderscheiden. De diagnose apathie werd gesteld aan de hand van de *Diagnostic Criteria for Apathy in Alzheimer's Disease and other Neuropsychiatric Disorders* (DCA). De AES-10 en NPIa correleerden significant en matig met elkaar (r_s = 0,62), wat impliceert dat hun congruente validiteit voldoende was. De AES-10 correleerde zwak (r_s = 0,27) en de NPIa matig (r_s = 0,46) met de CSDD, hetgeen impliceert dat de AES-10 de voorkeur geniet wat discriminante validiteit betreft. Er is geen significant verschil gevonden tussen beide instrumenten in hun vermogen om verpleeghuiscliënten met en zonder apathie te onderscheiden. De AES-10 scores hadden een hogere som van sensitiviteit en negatieve predictieve waarde dan de NPIa-scores, wat betekent dat de AES-10 de voorkeur heeft bij screeningsdoeleinden. Secundaire analyses lieten zien dat de AUCs significant waren bij cliënten zonder dementie (AES-10, AUC: 0,88; NPIa, AUC: 0,77), maar niet bij cliënten met dementie. Er werd geconcludeerd dat zowel de AES-10 als de NPIa gebruikt kunnen worden om cliënten met en zonder apathie te

onderscheiden in een heterogene groep cliënten met of zonder dementie, of bij cliënten zonder dementie.

Een nieuw model voor procesevaluatie helpt bij het verkrijgen van inzicht in de interne en externe validiteit van een interventie. Volgens ons model, gepresenteerd in hoofdstuk 6, bestaan de eerste-orde-procesgegevens (*first-order process data*) uit de *sample quality* (beschrijving en resultaten van de wervings- en randomisatieprocedures voor verpleeghuisafdelingen en hun cliënten) en de *intervention quality* (relevantie en uitvoerbaarheid van *een* zorgprogramma, en de mate waarin de interventie werd uitgevoerd zoals gepland). Deze gegevens zijn van belang voor inzicht in interne en externe validiteit. In tegenstelling tot tweede-orde-procesgegevens – die uitgevoerde implementatiestrategieën beschrijven, de blootstelling daaraan van deelnemende partijen en de belemmerende en bevorderende factoren – moeten de eerste-orde-procesgegevens geanalyseerd worden voorafgaand aan de statistische effectanalyses.

De evaluatie van sampling quality liet zien dat het noodzakelijk was om in de effectanalyses rekening te houden met het type afdeling. Het verloop en de resultaten van de informed consent (IC) procedure verschilden per type afdeling, wat impliceert dat het belangrijk is om rekening te houden met het type afdeling in verdere statistische analyses. Ondanks continue werving van nieuw opgenomen cliënten daalde het percentage cliënten met een IC op de psychogeriatrische (PG) afdelingen van 62% aan het begin van de studie tot 50% aan het einde van de studie. Het percentage IC's op de somatische afdelingen steeg juist geleidelijk van 52% tot 56% bij de vijfde meting en daalde tot 49% bij de zesde en tevens laatste meting. De IC-percentages waren te vergelijken met cijfers van andere studies in verpleeghuizen naar depressiezorg.

De evaluatie van intervention quality liet zien dat het noodzakelijk was om in de effectanalyses rekening te houden met de mate waarin elementen van het zorgprogramma uitgevoerd werden. Verpleeghuismedewerkers die betrokken waren bij de uitvoering van *Doen bij Depressie* waren tevreden met het zorgprogramma en rapporteerden dat het relevant en uitvoerbaar is in verpleeghuizen (externe validiteit). PG en somatische afdelingen verschilden in de mate waarin elementen van het zorgprogramma werden uitgevoerd. Verder bleek dat terwijl opsporings- procedures (signaleren, screenen en diagnosticeren) in 76% van de geïndiceerde gevallen werden uitgevoerd, behandelden de multidisciplinaire teams cliënten met een depressie of depressieve klachten slechts in 40% van de voorgeschreven gevallen. Het monitoren van de resultaten van de behandeling werd het minst uitgevoerd (18%). Omdat het zorgprogramma in de meeste gevallen niet in het geheel is uitgevoerd heeft dit

consequenties voor de interne validiteit. Er werd geconcludeerd dat in de effectanalyses niet alleen rekening gehouden moest worden met het type afdeling, maar ook met de mate waarin de elementen van het zorgprogramma werden uitgevoerd.

Het werken volgens *Doen bij Depressie* verlaagde de depressieprevalentie op de somatische afdelingen en verbeterde de kwaliteit van leven volgens cliënten met zowel somatische als PG problematiek. In hoofdstuk 7 worden de effecten van het werken volgens het zorgprogramma gepresenteerd voor depressieuitkomsten en de kwaliteit van leven. De depressieprevalentie, de primaire uitkomstmaat gemeten met de CSDD, daalde na het invoeren van het zorgprogramma in somatische afdelingen met 7,3% (absolute daling in procenten; 95% BI: -13,7 tot -0,9; Cohen's d: -0,6). De resultaten op de secundaire uitkomstmaten (de ernst van CSDD-symptomen en de kwaliteit van leven gemeten in interviews met cliënten) bevestigden dat het zorgprogramma effectief is op somatische afdelingen. Er waren ook aanwijzingen dat het zorgprogramma effectief kan zijn op PG afdelingen aangezien de kwaliteit van leven van cliënten op zowel somatische als PG afdelingen verbeterde. Na controle voor de mate waarin de afdelingen de voorgeschreven procedures van het zorgprogramma hadden uitgevoerd, bleek dat het interventie-effect vooral toegeschreven kan worden aan de uitgevoerde opsporingsprocedures (Cohen's d voor de primaire uitkomstmaat: -1,1). Behandelprotocollen werden minder vaak uitgevoerd en waren niet geassocieerd met depressie-uitkomsten. Er was geen effect op de zelfgerapporteerde depressiescores (de verkorte Geriatric Depression Scale met 8 items, GDS8). Deze scores ontbraken echter bij ruim de helft van de cliëntmetingen wat kan verklaren waarom op deze uitkomstmaat geen effect is gevonden. Wanneer de cliënten met cognitieve problemen waren meegenomen in de GDS8-analyses was er een indicatie dat de zelfgerapporteerde depressie daalde bij somatische cliënten. Zorgverleners beoordeelden de kwaliteit van leven van cliënten lager in de interventieconditie. De scores verbeterden echter bij het toepassen van de behandelprotocollen. Er werd geconcludeerd dat het zorgprogramma gebruikt kan worden om de depressieprevalentie te verlagen, vooral op somatische afdelingen. Omdat behandelprotocollen weinig toegepast werden en er ook aanwijzingen waren dat het zorgprogramma effectief kan zijn op PG afdelingen, is het essentieel om meer aandacht te besteden aan de verbetering van de implementatie van behandelprotocollen in verpleeghuizen.

Activerende behandelstrategieën voor cliënten met PG problematiek en psychotherapie voor cliënten met somatische problematiek kunnen effectief gebruikt worden voor het verminderen van apathie en van motivatieproblemen

bij depressie. In hoofdstuk 8 wordt gefocust op de effecten van het zorgprogramma op apathie en op twee groepen van depressieve symptomen, namelijk motivationele en stemmingsgerelateerde symptomen. Het zorgprogramma *Doen bij Depressie* als geheel verminderde apathie bij cliënten van PG-afdelingen en de ernst van depressieve motivationele symptomen (vertraging, verlies van interesse, gebrek aan eetlust, gewichtsverlies en gebrek aan energie) op somatische afdelingen. Stemmingssymptomen (angst, verdriet, prikkelbaarheid, agitatie, dagschommelingen, gedachten aan zelfmoord, laag zelfvertrouwen en pessimisme) werden niet beïnvloed door het zorgprogramma. De activerende behandelstrategieën - een dagprogramma en een plezierige-activiteitenplan in behandelmodule 1 - leverden de grootste bijdrage aan het verminderen van apathie (Cohen's d: -0,73) en werden geassocieerd met een verlaging van motivatieproblemen (Cohen's d: -0,45) bij cliënten van PG-afdelingen. Op somatische afdelingen leverde psychotherapie (behandelmodule 2) de grootste bijdrage aan het verminderen van motivationele symptomen (Cohen's d: -0,80), en werd deze geassocieerd met verminderde apathie (Cohen's d: -0,42). Een medicamenteuze behandeling van depressie bleek in verband te staan met een toename van apathie op zowel de PG- als de somatische afdelingen. Er werd geconcludeerd dat het zorgprogramma gebruikt kan worden voor het verminderen van apathie, en dat vooral activerende strategieën effectief gebruikt kunnen worden bij cliënten met PG-problematiek en psychotherapie bij cliënten met somatische problematiek. Met betrekking tot depressieve symptomen is het zorgprogramma vooral effectief gebleken voor motivationele symptomen. Meer onderzoek is nodig ten aanzien van medicamenteuze behandeling van depressie waarbij apathie niet verergert. Verder is er meer onderzoek nodig naar behandelstrategieën die effectief toegepast kunnen worden voor depressieve stemmingssymptomen.

Belangrijke conclusies

In hoofdstuk 9 worden de belangrijkste conclusies van dit proefschrift, methodologische en theoretische overwegingen besproken. Verder wordt ingegaan op de mogelijke implicaties voor toekomstig onderzoek en de praktijk.

Ten eerste kunnen de NORD, de CSDD en de AES-10 gebruikt worden voor screeningsdoeleinden in de verpleeghuispopulatie, maar kunnen deze niet een diagnostisch instrument vervangen.

Ten tweede onderstreept dit proefschrift het belang van het analyseren van een deel van de procesgegevens voorafgaand aan de statistische analyses. Dit is vooral van belang voor de zogenaamde pragmatic trials waarin effecten van een interventie onderzocht worden

in de dagelijkse praktijk en waarbij de condities niet strikt gestandaardiseerd (kunnen) worden. De procesevaluatie liet niet alleen zien dat het zorgprogramma relevant en uitvoerbaar werd gevonden door betrokken medewerkers, maar ook dat verdere analyses zinvol waren en verfijnd konden worden.

De belangrijkste bevinding van dit proefschrift is dat het *Doen bij Depressie* zorgprogramma ingezet kan worden voor het verlagen van de depressieprevalentie, voor de verbetering van de kwaliteit van leven van cliënten en het verlagen van apathie. Het *Doen bij Depressie*-onderzoek is een pragmatische trial die niet gericht was op een specifieke behandelinterventie, maar op het geheel van procedures en trajecten voor het opsporen en diagnosticeren van depressies en voor het behandelen daarvan in multidisciplinair verband. Het onderzoek naar *Doen bij Depressie* is uniek, omdat er verder geen studies bekend zijn in verpleeghuizen naar de effecten van structurele procedures voor het opsporen en screenen van depressies en depressieve klachten. Het heeft aangetoond dat het routinematig screenen op depressieve klachten, en het betrekken van afdelingsmedewerkers daarin, uitermate belangrijk is voor de verbetering van de depressiezorg. Helaas heeft de studie ook laten zien dat het invoeren van procedures voor het verbeteren van depressiezorg uitdagend blijft in verpleeghuizen. Hoewel de zorgmedewerkers voor bijna alle cliënten screeningslijsten hadden ingevuld, zijn voor een deel van de cliënten geen verdere stappen ondernomen in de opsporingselementen van *Doen bij Depressie*. Ook werden niet alle depressieve cliënten behandeld met psychosociale interventies. Dit werd deels toegeschreven aan de huidige economische uitdagingen - de bezuinigingen in de zorg en het personeelstekort. Toch is het ook belangrijk te beseffen dat de verbetering van depressiezorg niet volledig afhankelijk is van financiële investeringen. Sommige componenten van het zorgprogramma vragen bijvoorbeeld om relatief weinig tijd, zoals de eerste screeningsstap met behulp van de NORD. Voor andere componenten zal de tijdsinvestering steeds verder afnemen als gevolg van groeiende deskundigheid van personeel. Het gaat niet om wat we niet kunnen doen, maar om wat we juist wel kunnen doen om emotioneel lijden van cliënten te verminderen met de middelen die we hebben.

We kunnen Doen bij Depressie!

Appendix A1

Act in case of Depression program

A1.1 Program phases and pathways

A1.2 Choosing a cut-off score

A1.3 Abstract of AiD treatment modules

A1.1 Program phases and pathways

Act in case of Depression prescribes pathways for structural depression assessment, multidisciplinary treatment, and monitoring of treatment effects. Figure A1.0 represents the 5 phases of the program. Figure A1.1 provides a graphical representation of the pathways.

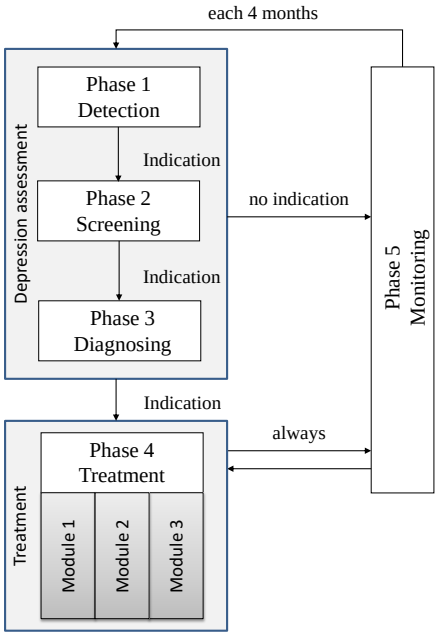


Figure A1.0 Graphical representation of the AiD care program

Depression assessment (phases 1, 2, 3)

Depression assessment contains 3 elements:

Detection (phase 1): A screening procedure executed by the nursing staff using a short observer-rated scale, Nijmegen Observer-Rated Depression scale (NORD),[1] cut-off score >1. See Section A1.2 for proposed cut-off scores. To reduce false negatives, psychologists may use other information for starting the second screening step (Screening).

Screening (phase 2): An extensive screening by the NH-unit psychologist using an interview-based instrument for the resident (Geriatric Depression Scale short version with 8 items [GDS-8][2], cut-off score>2) or for the caregiver if the resident has dementia or cannot be interviewed (Cornell Scale for Depression in Dementia [CSDD][3], cut-off score>7). See Section A1.2 for proposed cut-off scores. Psychologists and physicians may start the phase Diagnosing based on other additional information.

Diagnosing (phase 3). A diagnostic procedure by the psychologist and the elderly care physician using the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, revised (DSM-IV-TR)[4] in residents without dementia, and Provisional Diagnostic Criteria for Depression of Alzheimer's Disease (PDCdA)[5] in residents with dementia.

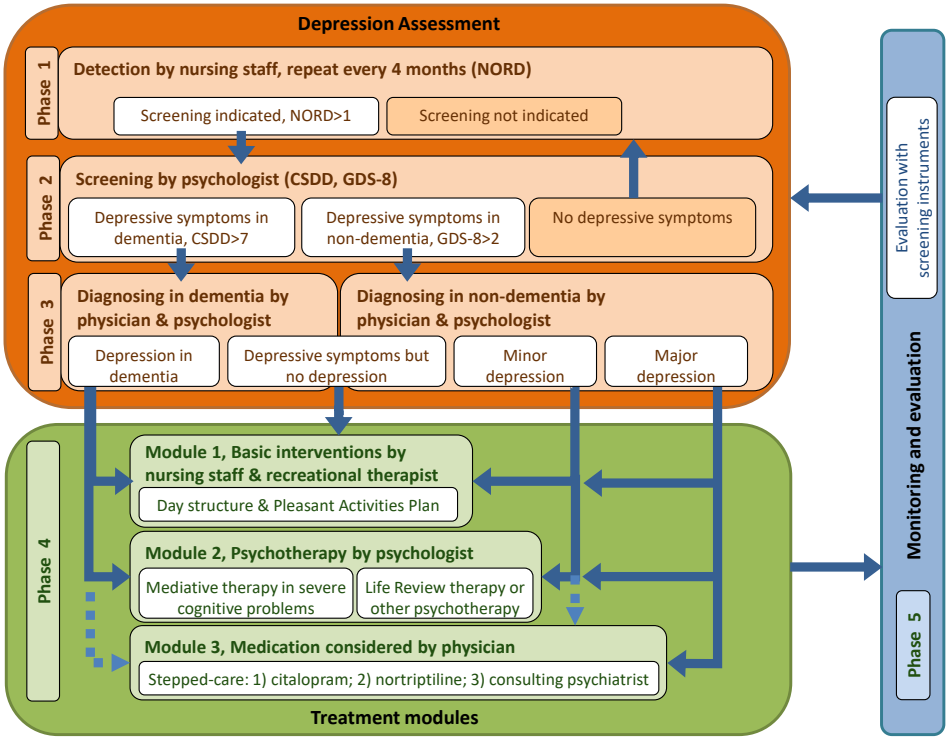


Figure A1.1 Act in case of Depression pathways

Light blue broken line: not prescribed but may be considered if symptoms are severe, or when a psychosocial treatment, i.e. module 1 and module 2, was not effective.

Treatment (phase 4)

The multidisciplinary team provide up to three treatment modules. The psychologist and physician can diverge from the protocols but should provide psychotherapy and consider a pharmacological treatment in accordance to the pathways. Abstracts of treatment protocols can be found in Appendix Section A1.3. Environmental and behavioral activation strategies (Module 1) are advised for residents with depressive symptoms or depression. Protocols of AiD Module 1 provide guidelines for psycho-education, a pleasant activities plan[6,7] and a day structure program. In case of depression, psychotherapy should be provided complementary to Module 1. Protocols of Module 2 provide guidelines for Life

Review therapy[8] by the unit psychologist (the first choice) or mediative therapy if communication is not possible due to language or cognitive problems. Module 3 includes the use of antidepressants complementary to modules 1 and 2, especially if depression is severe.

Monitoring (phase 5)

Monitoring prescribes the evaluation of the treatment: Screening is performed if the resident receives treatment. Residents without a depression and without depressive symptoms will be evaluated in phase 1 Detection.

A1.2 Choosing a cut-off score

Definitions

Sensitivity is the proportion of residents with depression who are identified as such. **Specificity** is the proportion of residents without depression who are identified as such. **Positive Predictive Value** (PPV) is the proportion of residents identified as depressed (positive test results) who are correctly diagnosed (true positives).

Negative Predictive Value (NPV) is the proportion of residents identified as not depressed (negative test results) who are correctly diagnosed (true negatives).

Positive Likelihood Ratio (LR+) is the probability of a resident with depression testing positive in relation to the probability of a resident without depression testing positive.

Negative Likelihood Ratio (LR-) is the probability of a resident without depression testing negative in relation to the probability of a resident with depression testing negative.

	Depression	No depression	
Positive test result (higher than cut-off score)	True Positive (TP)	False Positive (FP)	PPV = TP / (TP + FP)
Negative test result (less or equal than cut-off score)	False Negative (FN)	True Negative (TN)	NPV = TN / (TN + FN)
	Sensitivity = TP / (TP + FN)	Specificity = TN / (FP + TN)	

Positive Likelihood Ratio (LR+) = sensitivity / (1 – specificity)
Negative Likelihood Ratio (LR-) = specificity / (1 - sensitivity)

Which score to choose?

Ruling out condition	Screening for condition	Ruling in condition
⟷	⟷	⟷
Maximum (Sensitivity + NPV)	Optimal score?	Maximum (Specificity + PPV)

Choosing a cut-off score should, first, depend on the aims. An optimal cut-off score for ruling in a condition (confirming the diagnosis) will differ from an optimal cut-off score for ruling out a condition. Second, expected advantages (gains) and disadvantages (costs) of correct and incorrect decisions should be taken into account.

Ruling out: A test that provides a maximum of sensitivity (N of true positives is maximized) and NPV (N of false negatives is minimized) can be used to rule out a condition. Such a test can be used in the first screening phase of a two-step screening procedure.

Ruling in: A test that provides a maximum of specificity (maximum of non-depression is identified) and PPV (N of false positives is minimized) can be used to rule in a condition. Such a test can be used in the second phase of a two-step screening procedure.

Screening: A score between the cut-offs for ruling out and ruling in a condition can be used for screening purposes depending on the costs (benefits and losses) of incorrect decisions. Such a score can be used in a one-step or two-step screening procedure.

Distinguishing residents in two groups: A score that maximizes sensitivity and specificity can be used to distinguish residents with depression from those without depression. Such an optimal score can be used in clinical practice in case of equal costs of false positive and false negative decisions. A score with higher sensitivity is preferred when the costs of false negatives are high (e.g. suicidal risk in depression), and a score with higher specificity is preferred when costs of false positives are high (e.g. serious treatment side effects).

Using pretest probability: Likelihood ratios can be used to combine clinical suspicion, i.e. pretest probability, with a test result. A test with an LR+ greater than 10, or an LR- less than 0.1 is considered to have the potential to alter clinical decisions, i.e. to, respectively, rule in and rule out depression; and a test with an LR+ between 5 and 10, or an LR-between 0.1 and 0.2 as useful for additional information on screening for depression.[9]

See Table A1.1 for proposed cut-off scores of the scales validated in this thesis.

Table A1.1 Cut-off scores

	Ruling out condition, Score or range	'Optimal cut-off' [Sensitivity:Specificity]	Ruling in condition Score or range
Depression, NORD			
- in dementia	0*	> 2 [79:77]	5
- in non-dementia	0 - 1	> 1 [100:69]	n.a.
Depression, CSDD			
- in dementia	0 - 5	> 6 [94:49]	n.a.
Depression, MADRS			
- in dementia	0 - 7	> 13 [78:66]	n.a.
Apathy, AES-10			
- in dementia	10 - 16	> 29 [88:54]	n.a.
- in non-dementia	10 - 21	> 21 [100:75]	n.a.
Apathy, NPIa			
- in dementia	n.a.	> 4 [51:76]	n.a.
- in non-dementia	n.a.	> 0 [67:89]	n.a.

'Optimal cut-off': a cut-off score with a maximum of sensitivity and specificity. This score is optimal when the costs of false negative and false positive decisions are equal.
0*: this NORD score of '0' in dementia showed a sensitivity of 95%
Abbreviations, n.a.: not applicable; NORD: Nijmegen Observer-Rated Depression scale; CSDD: Cornell Scale for Depression in Dementia; MADRS: Montgomery-Åsberg Depression Rating Scale; AES-10: ten-item abbreviated Apathy Evaluation Scale (AES-10); NPIa: Neuropsychiatric Inventory apathy subscale (NPIa)

A1.3 Abstract of AiD treatment modules

The full texts in Dutch can be requested at www.uko-n.nl.

1. Treatment Module 1

For whom: Nursing home residents with depressive symptoms (with or without an official diagnosis of depression).

Module elements: The module consists of a day structure program (DSP) and a pleasant activities plan (PAP).[6,7] The DSP is mainly aimed at structuring daily activities and avoiding disruptions in circadian rhythm. The PAP is aimed at involving the client in activities that are experienced as pleasant, and avoiding activities that are experienced as unpleasant.

Implementers: If there is a recreational therapist available in the unit, she or he develops a PAP. If there is no recreational therapist, a primary professional caregiver develops both a PAP and a DSP. The unit psychologist supervises the module development. In case module 2 is prescribed by the AiD algorithm, the psychologist advises the nursing team proactively about the module development. Nursing staff and a recreational therapist apply the module in daily care.

Content: The developed DSP and PAP are goal-directed and solution-based, i.e. NH staff sets initial goals (such as "the client socializes with other residents") and researches and describes possible solutions. The developers and implementers use information about the resident from the resident's personal file and interviews with the resident, and other persons including family or relatives who may provide information about the resident.

The staff may use templates and examples of DSP and PAP provided in the texts of the AiD care program, or they may use other available tools in their organization. In all cases, the implementers monitor whether the AiD quality criteria are met for the contents of the module.

Quality criteria for the elements of Module 1

The implementers monitor whether the elements meet the following criteria:

- **Day structure program:** For stabilizing the circadian rhythm pattern, the regularity of waking-up time, eating and sleeping is enhanced; the resident receives as much daylight as possible; if possible, physical activities are encouraged; social contacts are stimulated; a healthy food pattern is stimulated (e.g. food rich in fruit and vegetables); the client is stimulated to make his or her own choices in activities and to be as independent as possible with regard to activities of daily living. A DSP provides information on the planning of, and the persons responsible for, the evaluation of its implementation and content.

- **Pleasant activities plan:** The resident receives daylight and sunlight; if possible, indoor and outdoor physical activities are introduced; regular social contacts are stimulated; activities are introduced that are in line with the resident's personal interests and hobbies (past and present interests and hobbies are explored); the activities are meaningful, pleasant and feasible for the resident; activities that are unpleasant are avoided. A PAP provides information on the planning of, and the persons responsible for, the evaluation of its implementation and content.

The AiD care program does not provide quantitative recommendations for individual quality criteria. The use of national guidelines (if available) or obtaining professional advice from the (para)medical staff is advised.

2. Treatment Module 2

For whom: Nursing home residents with minor depression, major depression or depression in dementia. The psychologist may decide – when needed – to introduce module 2 to residents with depressive symptoms but without a depression diagnosis.

Protocols: The module makes a distinction between residents with and without cognitive problems. The first protocol is for a mediative therapy complemented with individual sessions for residents with (severe) cognitive or language problems that impede 'talk' psychotherapy. The second protocol is for a type of life-review therapy for residents without cognitive problems.

Implementer: A nursing home psychologist implements the therapy complementary to module 1 (basic psychosocial intervention provided by the nursing staff and a recreational therapist).

Content: Therapy starts with an individual session with a psychologist to explore whether a 'talk therapy' may be used. The content of the treatment which follows depends on the protocol chosen. The psychologist may decide – if appropriate – to provide an evidence-based psychosocial intervention not described in one of the two protocols.

2.1 Module 2 for residents with (severe) cognitive problems

This variant of module 2 comprises two elements: 1) five individual sessions with the resident, and 2) a mediative intervention.

- **Individual sessions:** If communication with the resident is possible, the first individual contact with the resident is aimed at exploring the applicability of a 'talk therapy'. If such therapy cannot be applied, but verbal or non-verbal communication is still possible, four additional contacts with the resident should be planned (once every two weeks). The aims are: obtaining additional information about the client and monitoring depressive symptoms and the

treatment effects; contributing to the client's feeling of well-being by providing individual attention; exploring the possible objectives of the mediative intervention and obtaining additional input for the fine-tuning of treatment module 1. An individual session lasts 20 to 30 minutes and ends with a pleasant activity.

- **Mediative interventions:** This element is implemented alongside the individual sessions with the client. The psychologist aims to treat the resident by intervening in the client's (social) environment. This can be realized in the nursing home by educating and consulting the nursing staff. Through analyzing and interpreting the meaning of the client's behavior, the staff receives tools to cope with the client's behavior, to reframe it and to react in a way that benefits the client. In total, the psychologist plans five consulting sessions with the nursing staff. In the first session, the psychologist explains the therapy concepts and prioritizes which behavior will be focused on in the upcoming days. Staff interactions that are useful and those that are to be avoided are discussed. In the following sessions, the staff and the psychologist discuss and evaluate the functionality of the client's behavior and mood, and of the staff reactions. They discuss (new) possible mediative interventions and choose the strategy for the upcoming period. The psychologist uses the individual sessions with the client as input for the fine-tuning of the mediative interventions.

2.2 Module 2 for residents without cognitive problems

'Dierbare Herinneringen' therapy *(DHT, 'precious memories therapy'*)[10] is a form of life review therapy adapted for the nursing home population. According to AiD, DHT is the first-choice psychotherapy provided to residents with depression. However, the psychologist may decide to provide another evidence-based therapy if appropriate.

DHT is based on the concept that retrieving, reaffirming and validating positive memories in the context of present day life may help treat depression. A depressed person may have overgeneralised autobiographical memories, whereas training to retrieve specific positive memories will improve his or her mood (for theoretical discussion see Haber[11]). The therapy is a form of cognitive training, and it can be less effective if a person cannot be trained due to cognitive problems. It comprises at least 5 individual sessions of 45 minutes each, provided once a week or once every two weeks. The first session is used to explore whether the training is suitable, and which life periods should receive more or less focus in subsequent sessions that cover childhood until the age of 12 years (session 2), adolescence until the age of 18 years (session 3), adulthood (one or two sessions), and the whole life (last session). Every session starts by repeating and explaining the aims and the theoretical concepts of DHT. Homework is discussed, the life period is selected, after which

the psychologist starts the training. A template is used with questions such as “can you remember and describe your favorite toy to me?” (session about childhood, the client will then be stimulated to describe the toy and retrieve specific positive memories about events with the toy). In total, the client will be stimulated to retrieve 8 to 10 specific positive memories in one session. The psychologist uses positive reinforcements if the client succeeds in retrieving a memory that is positive and specific. The session ends with a short evaluation and discussion of a homework assignment. The assignment can be used to further train the client to retrieve specific positive memories. Additional material such as personal photos can be used.

3. Treatment Module 3

For whom: The physician considers a pharmacological treatment in major depression. For residents with minor depression or depression in dementia, module 3 may be considered if depressive symptoms are severe, or when a multidisciplinary psychosocial treatment (module 1 and module 2) was not effective or showed weak improvement in depressive symptoms.

Two protocols: The module has two protocols based on national guidelines for treatment of depression in elderly:[12] one for residents without dementia, and a separate protocol for patients with dementia, which is in line with the main protocol but emphasizes the need of psychosocial treatment prior to pharmacological treatment.

Implementer: an elderly care physician implements the therapy complementary to module 1 (basic psychosocial interventions) and module 2 (psychotherapy).

Content: First-choice medication is Citalopram (10 mg increasing to 20 mg a week; 40 mg in case of partial effect). Treatment effects are evaluated after four weeks. When Citalopram proves to be ineffective, Nortriptyline is the second-choice (starting with 10-25 mg, increasing to 25-75 mg in 2-4 days; blood levels are to be controlled to avoid toxicity and to optimize response). If Nortriptyline is contra-indicated, Venlafaxine is the second-choice treatment (>150 mg).

The protocols indicate that a psychiatrist is to be consulted in case of psychotic symptoms, suicidal attempts, or food refusal. This is also the case if prescribed medication remains ineffective and/or when lithium addition is considered. Only a psychiatrist may indicate a more invasive therapy such as ECT (electroconvulsive therapy).

Literature

1. Leontjevas R, Gerritsen DL, Vernooij-Dassen MJFJ, et al. The Nijmegen Observer-Rated Depression Scale for Detection of Depression in Nursing Home Residents. *Int J Geriatr Psychiatr*. In press.

2. Jongenelis K, Gerritsen DL, Pot AM, et al. Construction and validation of a patient- and user-friendly nursing home version of the Geriatric Depression Scale. *Int J Geriatr Psychiatry* 2007;22:837-842.
3. Alexopoulos GS, Abrams RC, Young RC, et al. Cornell Scale for Depression in Dementia. *Biol Psychiat* 1988;23:271-284.
4. APA. Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, Text Revision. Washington, DC: American Psychiatric Association, 2000.
5. Olin JT, Schneider LS, Katz IR, et al. Provisional diagnostic criteria for depression of Alzheimer disease. *Am J Geriatr Psychiatry* 2002;10:125-128.
6. Teri L, Logsdon RG, Uomoto J, et al. Behavioral treatment of depression in dementia patients: a controlled clinical trial. *J Gerontol B Psychol Sci Soc Sci* 1997;52:P159-166.
7. Verkaik R, Francke AL, van Meijel B, et al. The effects of a nursing guideline on depression in psychogeriatric nursing home residents with dementia. *Int J Geriatr Psychiatry* 2010.
8. Bohlmeijer E, Steunenberg B, Leontjevas R, et al. *Dierbare herinneringen Protocol voor individuele life-review therapie*. Enschede: Universiteit Twente, 2010.
9. Fischer JE, Bachmann LM, Jaeschke R. 2003. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003; 29: 1043–1051.
10. Bohlmeijer E, Steunenberg B, Leontjevas R, et al. In Ducth: [Dierbare Herinneringen: Protocol voor Individuele life-review therapie gebaseerd op autobiografische oefening]. Enschede: Universiteit Twente, 2010.
11. Haber D. Life review: implementation, theory, research, and therapy. *Int J Aging Hum Dev* 2006;63:153-171.
12. Trimbos-instituut. In Dutch: [Addendum ouderen bij de multidisciplinaire richtlijn depressie]. Utrecht: Trimbos-instituut, 2008.

Appendix A2

Dankwoord

Dankwoord

Dit proefschrift is een product van ruim duizend mensen. Ik kon de opgenomen artikelen met mijn co-auteurs schrijven omdat bijna 800 verpleeghuisbewoners en hun naasten wilden meewerken aan het onderzoek. Omdat ruim 900 zorgmedewerkers, activiteitenbegeleiders, psychologen en artsen van de 33 deelnemende afdelingen hun tijd vrijmaakten voor de interviews, het lezen van programmateksten, volgen van cursussen, uitvoeren van het zorgprogramma op de afdelingen en het invullen van diverse formulieren. Omdat leidinggevenden van de afdelingen mensen hebben vrijgemaakt tijdens de interviews en het onderzoek op de werkvloer hebben gecoördineerd; omdat … Omdat deze mensen aan een onderzoek voor de verbetering van de depressiezorg in verpleeghuizen wilden meewerken en dit met bevlogenheid hebben gedaan. Beste mensen, bedankt! Bedankt dat jullie dit onderzoek mogelijk hebben gemaakt!

Ik heb het ontzettend getroffen met mijn onderzoeksteam. Behalve de kern van het team waren er 25 interviewers die, verspreid over twee jaar, interviews hebben afgenomen in 22 verpleeghuizen. Amber, Cissy, Danielle, Denise, Edwin, Esther, Gerda, Harry, Helen, Ilse, Joke, Kim, Lieke, Lily, Maaike, Marianne, Marijke, Michelle, Patrick, Renée, Rianda, Ribanna, Saskia, Stéphanie en Sule - zonder jullie zou ik geen gegevens hebben gehad om dit proefschrift te schrijven. Dank jullie wel!

Annelies, jij bent als assistente bijna tegelijk met mij gestart. Bedankt voor je hulp en dat je je met veel energie ingezet hebt voor het project! Ellen, je was de tweede engel die me te hulp schoot. Ik voel me bevoorrecht om een toekomstige succesvolle advocaat of rechter - wat je vast ooit zult worden - te hebben ontmoet. Zonder jouw doordachte werkplanningen, jouw energie en inzet zou het project minder soepel zijn verlopen. Dit zou ook het geval zijn geweest als jij, Alex, er niet was geweest. Nog tijdens je studie aan de Open Universiteit raakte je betrokken bij het project en je bent gebleven om ons team te versterken als onderzoeksmedewerkster. Alex, zonder jou zou ik nog lang niet klaar zijn geweest met dit proefschrift. Je hebt me enorm geholpen en ik heb in alle opzichten enorm veel steun aan je gehad. Bedankt! Ellen en Alex, bedankt dat jullie mijn paranimfen wilden zijn tijdens de verdediging van dit proefschrift. Een mooie afsluiting van onze vruchtvolle samenwerking.

Behalve Ellen en Alex waren er nog Ans, Jessika, Noortje en Janine, die tijdelijk of met tussenpozen tot het team behoorden. Bedankt voor jullie inzet, de vele gezellige momenten en de gezamenlijke lunches. Ik wens jullie allemaal veel succes met jullie verdere carrières als psychologen!

Arjanne, ook jij bent een engel. Zonder je hulp in de eindfase zou ik niet in september hebben kunnen promoveren. Dankzij jou kunnen meerdere wetenschappelijke artikelen geschreven worden die als input kunnen dienen voor een betere depressiezorg. Bedankt!

Raymond, het projectteam kwam naast jouw kamer te zitten en ik hoop dat je niet teveel last van ons hebt gehad. We hebben vaak van je geweldige lach genoten. Ik heb veel aan je begeleiding gehad tijdens het onderzoek en het schrijven van dit proefschrift en heb veel respect voor jouw positief motiverende manier van feedback geven. Ik wens van harte dat je je onderzoeksideeën blijft verwezenlijken en je nog veel promovendi kunt enthousiasmeren en stimuleren!

Myrra, dankzij jou ben ik veel gaan lezen over procesevaluatie en is hier een artikel over in mijn proefschrift opgenomen. Bedankt voor je bijdrage aan onze discussies!

Martin, fijn dat je mijn co-promotor bent geweest. Door op een gestructureerde manier discussies te voeren creëerde je rust in mijn hoofd en droeg je bij aan een gevoel van zekerheid. In een project zoals *Doen bij Depressie* was iemand zoals jij een must. Bedankt voor je hulp en steun!

Steven, enorm veel dank voor je hulp, de uitleg die je gaf en de tijd die je voor me nam als ik langs kwam met vragen over statistiek. Ik was een lastige “klant” want ik had steeds nieuwe ideeën voor aanvullende berekeningen. Mijn petje af voor hoe je met je “klanten” omgaat.

Els, wat bofte ik en het *Doen bij Depressie* onderzoeksteam met je steun, met je bijdrage aan de ontwikkeling van de cursus en je meedenken aan de uitvoering van het project. Fijn dat ik altijd met mijn vragen bij je terecht kon. Bedankt!

Ronald, Sandra en Johan, bedankt voor jullie praktische tips voor de productie en het drukwerk!

En nu… voordat ik nog iemand noem die bij het project was betrokken wil ik nog enkele vrienden bedanken: Lena, Iryna, François, Marina, Galina, Julia, Yulija, Nastja, Frank, Vincent – bedankt dat jullie er voor me waren en me bleven wijzen op het feit dat er meer is dan *Doen bij Depressie*! Frank en Vincent, bedankt voor jullie voorbeeldrol hoe je succesvol kunt promoveren en toch nog andere leuke dingen ernaast kunt blijven doen. Robert, bedankt dat je er was en dat je er bent, dat je me steunde en steunt. Er zouden veel meer taalfouten in dit proefschrift te vinden zijn zonder jou. Bedankt! Ook Scott bedankt voor je redactiewerk! Fijn dat je steeds extra tijd nam om het correctiewerk uitvoerig te bespreken. Christ, bedankt dat je me steunde bij het nemen van de beslissing om aan een promotietraject te beginnen.

Mama - Vera – zonder je vele hulp zou het promotietraject zwaarder zijn geweest. Bedankt dat je er altijd voor me bent!

Mijn dankwoord wil ik afsluiten met in het zonnetje zetten van iemand die haar energie, “bloed, zweet en tranen” aan het project heeft besteed en wie nummer één is geweest in *Doen bij Depressie*. Debby, je hebt niet alleen ontzettend veel gedaan voor het project maar ook voor mijn persoonlijke groei. Inhoudelijk heb je een mooi project laten opbloeien. Maar ook verloor je nooit het menselijke aspect in ons werk uit het oog. Al hadden we het enorm druk en was er sprake van flinke stress, je plaatste nieuwsbrieven en persoonlijke communicatie met afdelingen hoog op agenda. Je hebt altijd goed gezorgd voor iedereen in het onderzoeksteam en hebt ons emotioneel gesteund als het nodig was. Je vergeet niet waar het uiteindelijk om moet gaan in de wetenschap. Namelijk om mensen, of dit nu de onderzoekers zelf of de deelnemers zijn. Integriteit, ethiek en respect zijn geen losse woorden voor je. Je collega’s en je meerderen kunnen trots zijn op een senior onderzoeker zoals jij. **Bedankt, Debby!**

Appendix A3

Proefschriftenlijst Radboud Alzheimer Centrum

Van Horssen, J. (2005). *Heparan sulfate proteoglycans and vascular pathology in Alzheimer's disease*

Wilhelmus, M.M.M. (2006). *Small heat shock proteins and apolipoprotein E in Alzheimer's disease*

Zuidema, S.U. (2008). *Neuropsychiatric symptoms in Dutch nursing home patients with dementia*

Graff, M.J.L. (2008). *Effectiveness and efficiency of community occupational therapy for older people with dementia and the caregivers*

Claassen, J.A.H.R. (2008). *Cerebral hemodynamics in aging: the interplay between blood pressure, cerebral perfusion, and dementia*

Dado- Van Beek, H.E.A. (2010). *The regulation of cerebral perfusion in patients with Alzheimer's disease*

De Jong, D. (2010). *Anti-inflammatory therapy and cerebrospinal fluid diagnosis in Alzheimer's disease*

Persoon, J.W.B. (2010). *Development and validation of the Nurses' Observation Scale for Cognitive Abilities – NOSCA*

Meulenbroek, O.V. (2010). *Neural correlates of episodic memory in healthy aging and Alzheimer's disease*

Timmer, N.M. (2011). *The interaction of heparan sulfate proteoglycans with the amyloid β protein*

Schölzel-Dorenbos, C.J.M. (2011). *Quality of life in dementia: From concept to practice*

Bruinsma, I.B. (2011). *Amyloidogenic proteins in Alzheimer's disease and Parkinson's disease: interaction with chaperones and inflammation*

Perry, M. (2011). *Development and evaluation of a Dementia Training Programme for Primary care*

Derksen, E.W.C. (2011). *Diagnostic disclosure: a basic intervention in dementia care*

Wetzels, R.B. (2011). *Neuropsychiatric symptoms in institutionalized residents with dementia: Course and interplay with cognition,quality of life and psychotropic drug use.*

Voight-Radloff, S. (2012). *Cross-national transfer of community occupational therapy in dementia.*

Spies, P.E. (2012). *The reflection of Alzheimer disease in CSF.*

Joosten-Weyn Banning, E.W.A. (2012). *Learning to live with Mild Cognitive Impairment: development and evaluation of a psychological intervention for patients with Mild Cognitive Impairment and their significant others.*

Vasse, E. (2012). *A stepwise process for developing and implementing quality indicators to improve psychosocial dementia care in European countries.*

Slats, D. (2012). *CSF biomarkers of Alzheimer's disease; serial sampling analysis and the study of circadian rhythmicity.*

Leontjevas, R (2012). *Act in case of Depression! Validation and effectiveness of a multidisciplinary depression care program in nursing homes*