Development of an International Schedule for the Assessment and Staging of Care for Dementia

Maya Semrau[a,*], Alistair Burns[b], Slavica Djukic-Dejanovic[c], Defne Eraslan[d], Changsu Han[e],
Dusica Lecic-Tosevski[f], Antonio Lobo[g], Adriana Mihai[h], Julie Morris[i], Claudia Palumbo[j],
Philippe Robert[k], Gerthild Stiens[l], Gabriela Stoppe[m], Umberto Volpe[n], Marcel Olde Rikkert[o]
and Norman Sartorius[p] on behalf of the International Dementia Alliance (IDEAL) study group
[a]*King's College London, Institute of Psychiatry, London, UK*
[b]*University of Manchester and Manchester Academic Health Science Centre (MAHSC), Manchester, UK*
[c]*School of Medicine, University of Kragujevac, Kragujevac, Serbia*
[d]*Istanbul Psikiyatri Enstitüsü, Istanbul, Turkey*
[e]*Department of Psychiatry, Korea University College of Medicine, Seoul, Korea*
[f]*School of Medicine, University of Belgrade, Serbian Academy of Sciences and Arts, Belgrade, Serbia*
[g]*Department of Psychiatry, Universidad Zaragoza and Instituto de Investigación Sanitaria de Aragón, Centro de
Investigación Biomédica en Red de Salud Mental (CIBERSAM), Ministry of Science and Innovation, Madrid, Spain*
[h]*Department of Psychiatry, University of Medicine and Pharmacy, Tg Mures, Romania*
[i]*University Hospital of South Manchester, Manchester, United Kingdom*
[j]*Department of Medical Basic Sciences, Neuroscience and Sense Organs, University of Bari, Bari, Italy*
[k]*Memory CenterCHU – CoBTeK Research, University of Nice Sophia Antipolis, Nice, France*
[l]*Center for Old Age Psychiatry, LVR Hospitals Bonn, Bonn, Germany*
[m]*MentAge Consulting – Practice-Research, Basel, Switzerland*
[n]*Department of Psychiatry, University of Naples SUN, Naples, Italy*
[o]*Radboud Alzheimer Centre/Department Geriatrics, Radboud University Medical Centre, Nijmegen,
The Netherlands*
[p]*Association for the Improvement of Mental Health Programmes, Geneva, Switzerland*

Abstract.
Background: A reliable and valid global staging scale has been lacking within dementia care.
Objective: To develop an easy-to-use multi-dimensional clinical staging schedule for dementia.
Methods: The schedule was developed through: i) Two series of focus groups (40 and 48 participants, respectively) in Denmark, France, Germany, Netherlands, Spain, Switzerland, and UK with a multi-disciplinary group of professionals working within dementia care, to assess the need for a dementia-staging tool and to obtain suggestions on its design and characteristics; ii) A pilot-study over three rounds to test inter-rater reliability of the newly developed schedule using written case histories, with five members of the project's steering committee and 27 of their colleagues from Netherlands, France, and Spain as participants; and iii) A field-study to test the schedule's inter-rater reliability in clinical practice in France, Germany, Netherlands, Spain, Italy, Turkey, South Korea, Romania, and Serbia, which included 209 dementia patients and 217 of their caregivers as participants.

*Correspondence to: Dr. Maya Semrau, Institute of Psychiatry, King's College London, Box P029, De Crespigny Park, Denmark Hill, London SE5 8AF, UK. E-mail: maya.semrau@kcl.ac.uk.

Results: Focus group participants indicated a clear need for a culture-fair international dementia staging scale and reached consensus on face validity and content validity. Accordingly, the schedule has been composed of seven dimensions including behavioral, cognitive, physical, functional, social, and care aspects. Overall, the schedule showed adequate face validity, content validity, and inter-rater reliability; in the nine field-sites, intraclass correlation coefficients (ICCs; absolute agreement) for individual dimensions ranged between 0.38 and 1.0, with 84.4% of ICCs over 0.7. ICCs for total sum scores ranged between 0.89 and 0.99 in the nine field-sites.
Conclusion: The IDEAL schedule looks promising as tool for the clinical and social management of people with dementia globally, though further reliability and validity testing is needed.

Keywords: Dementia, mental status schedule, patient schedule, psychometrics, reliability, validity

INTRODUCTION

Over the last few decades, dementia has developed into a major challenge among older people worldwide, and it is likely to become an even larger issue in the future. While in 2010, 35.6 million people worldwide had dementia, it is estimated that in 2030, 65.7 million people will be affected, and that this number will have grown to 115.4 million by 2050 [1, 2]. This increasing prevalence of dementia, together with the associated burden accompanying the illness (on patients, families, and carers) lays high expectations and an increasing demand for care on health and social care systems, and calls for a consensus on the issue of access to diagnosis and care. In daily practice, the coordination of care is often lacking and the provision of professional care is mostly reactive or targeted at short-term goals, rather than being proactive, and based on the social circumstances of the patient [3], with few countries worldwide having national dementia plans, particularly countries outside of Europe and North America [1].

Staging models in which diseases or disorders are assessed according to different severity levels of the disease or disorder have been shown to be useful within cancer care [4], as well as for various mental disorders, such as bipolar disorder [5–9], schizophrenia [5, 7, 10–12], and depression [5, 11–14]. However, this work has not yet resulted in an evidence or consensus-based standard to improve and limit the large variability in prescription of treatments at different stages of these disorders. Although staging has also been applied for dementia in several ways [15, 16], this has not yet been extended to dementia care needs. Taking inspiration from staging care models for other disorders, it might be the case that within dementia care a more precise description of care at different severity levels of the disorder, together with a simplification and unification of interventions, and the communication of the results of care, would make it possible to provide better help and improve the international quality, homogeneity, and efficiency of dementia care (although this has

not been tested yet). To make a comprehensive coordinated care plan for dementia, it would be advantageous to have reliable information about the patient's condition across multiple domains (for example cognition or behavior), as well as about the perceived need for (professional) care.

However, a reliable and valid global planning scale, which could be used to guide the clinical and social care of people with dementia, is still lacking. A systematic review conducted by the International Dementia Alliance (IDEAL) group showed that in the last thirty years, twelve clinical dementia staging scales had been developed, of which the Clinical Dementia Rating (CDR) Scale was the best-evidenced one [17]. None of the twelve scales was complete in that it had been well validated, showed good reliability, was applicable in the entire course of dementia, and had been used or standardized for use in different cultures, i.e., was applicable across cultures [17]. The IDEAL group, which aims to develop an international standard of dementia care, was formerly named the European Dementia Consensus Network (EDCON), as the group first focused on developing consensus statements around controversial issues in dementia care, for example on access to care, genetic testing, and informed consent [18–21].

This paper describes the development of an internationally applicable schedule for dementia care by the IDEAL group. The approach used differs from others in two ways: i) A wide/comprehensive range of dimensions was used to define the different stages of dementia; and ii) the schedule was produced by an international group and intended for use in different countries.

Overview of IDEAL schedule

The IDEAL schedule consists of seven dimensions (see Supplementary Material): activities of daily living; physical health; cognitive functioning; behavioral and psychological symptoms; social support; non-professional care (which includes two sub-dimensions:

time spent on care by non-professional carer and carer distress); and professional care (which includes three sub-dimensions: total number of hours of professional care received, total number of hours of professional care needed, and type of dementia-related care needed). Each dimension is rated on a six-point scale of 0 to 5, with anchor points (for which there is an accompanying glossary of anchor point definitions) denoting the meaning of ratings. A sum score can also be calculated by adding up the individual scores of the seven dimensions; for the two dimensions with sub-dimensions, the average score of the sub-dimensions is calculated first before adding them to the total score. Total scores can range between 0 and 35. The IDEAL schedule may therefore give a preliminary indication of the overall level of care required (denoted by the sum score), as well as (more importantly) providing information about the different aspects of dementia symptomatology and care requirements (denoted by the individual dimensions), which in turn may give more detailed information about the type and level of care required.

MATERIALS AND METHODS

Design

The development of the IDEAL schedule was guided by a multi-disciplinary steering committee of international dementia experts, who had already addressed several other controversial issues relating to dementia [18–21]. Ethics approvals were obtained through the local ethics boards in each of the organizations through which data were collected. Three stages of scale development were employed in the development of the IDEAL schedule, according to the following procedures:

1. Two sets of focus groups were conducted with professionals working within dementia care. The aim of the focus groups was to assess the need for a new dementia-staging tool, and to obtain suggestions on the design, the necessary items, and characteristics of such a tool. This safeguarded face validity and partially also content validity.
2. Preliminary inter-rater reliability testing of the first draft of the schedule was carried out during three rounds of a pilot-study, using written case histories.
3. A large-scale field study was carried out, testing the inter-rater reliability of the IDEAL schedule when used to assess patients in clinical practice.

Ad 1. Focus groups

Sample

Forty expert participants from six European countries took part in the first round of focus groups (nine in Germany, seven in France, seven in the Netherlands, six in Denmark, six in Switzerland, and five in the UK), and 48 participants from six countries in the second round (thirteen in Spain, eleven in the Netherlands, eight in the UK, six in France, five in Denmark, and five in Switzerland). Focus group participants included relevant stakeholders who were involved with the care of people with dementia (i.e., professionals, caregivers, and representatives of patients) in participating countries; these included psychiatrists, psychologists, general practitioners, nurses, geriatricians, caregivers, and social workers as well as several other professionals working within dementia care.

Setting

Focus groups were conducted within the health care facilities of each participating European center.

Procedures

Both sets of focus groups were facilitated by members of the IDEAL steering committee in each participating center. Focus group discussions were conducted in the local language of each center; a summary of the major conclusions of the discussions were then provided by each of the focus group facilitators in English.

The dementia-staging tool was described for the participants of the focus groups as a global instrument that would allow to validly and reliably assess the severity of symptoms and related disease burden for the patient and caregiver, and the dementia-related need for health care services and informal care. During the first set of focus groups, the following questions were discussed:

- Is it helpful, or is there a need, to define stages of dementia?
- What, if any, instruments are used at present?
- Are these instruments fit for purpose and, if not, how can they be improved?
- What are the thresholds for a change from one stage to another?
- Would a change in threshold automatically lead to a change in care?
- What domains of staging would be useful, e.g. disease, disability, symptoms, nursing care?
- Would a multi-axial model that takes these into consideration be helpful?

- Is there a need for a brief SWOT analysis (strengths, weaknesses, opportunities, threats) of staging for dementia?
- Will staging be applicable to all types of dementia?
- What would an ideal instrument look like?

In some countries (e.g., Netherlands), peer-reviewed literature was drawn on and referred to in the discussions (e.g., [22–24]). A first draft of dimensions for inclusion into the draft IDEAL schedule, together with their anchor points, were developed by the IDEAL steering committee based on the first set of focus groups.

During the second set of focus groups, participants rated each of these draft dimensions in discussion using six specially-developed case vignettes, to assess their intelligibility, applicability, and content. A first draft of the IDEAL schedule was then developed based on these results through discussions and expert consensus within the IDEAL group's steering committee.

Ad 2. Pilot-study

Sample

The pilot-study was conducted by email first with five members of the IDEAL steering committee from five different European countries (France, Netherlands, Spain, Switzerland, and the UK), as well as twelve of their colleagues (i.e., dementia care professionals) in the Netherlands (round 1); then repeated with five members of the IDEAL group with a revised version of the draft IDEAL schedule (round 2); and finally carried out with fifteen further colleagues of IDEAL members in three European countries (seven in Spain, six in the Netherlands, and two in France) after the schedule had been revised even further (round 3).

Procedures

Ten specially-composed case histories of dementia patients (five male and five female, with varying degrees of symptom severity) were used by the expert participants of the pilot-study to make ratings on the draft IDEAL schedule. Primarily, English versions of all project materials were used, which were then translated locally by the researchers.

Ad 3. Field-study

Sample

Field-study participants included 209 dementia patients with varying degrees of cognitive decline, together with 217 of their caregivers, across nine countries (Nijmegen, Netherlands; Zaragoza, Spain; Nice, France; Bonn, Germany; Bari, Italy; Istanbul, Turkey; Seoul, South Korea; Tg Mures, Romania; and Belgrade/Kragujevak, Serbia). Patients' cognitive decline was approximated through use of the CDR (very mild to mild: CDR score of 0.5–1; moderate: CDR score of 2; or severe: CDR score of 3 or above), as this had been found to be the best-evidenced existing dementia scale [17], though possibly less so for very severe dementia.

Settings

Data collection for the field-study was conducted within the health care facilities of each participating center.

Procedures

Sample size calculations showed that a minimum of fourteen patients and their carers in each participating center would lead to an accuracy of the intra-class correlation coefficient (ICC) estimates corresponding to 95% confidence intervals of approximately ± 0.2. Interviews were conducted as part of patients' routine examinations; consecutive patients for diagnostic work-up for dementia in primary or secondary care were included into the study in each center, together with their caregivers. Patients were excluded from the study if there was no caregiver available, or if the contact between the caregiver and the patient was insufficient (i.e., less than once a week); if the patient did not speak the native language; if there was an unclear diagnosis after the diagnostic work-up; if a non-neurodegenerative disease was the cause of the cognitive disorder (for example brain tumor); or if the patient showed delirium.

During interviews, two raters (one interviewer and one silent observer/rater) made ratings (without consultation, blind to each other) for patients and their carers on the draft IDEAL schedule. Interviewers and silent raters rotated their role. Interviewers divided the interview time between the patient and the carer depending on the patient's symptom severity. There was no time restriction for interviews.

Interviews were semi-structured, and example questions were provided for each of the IDEAL schedule's dimensions as follows:

- How dependent is the person on others? Who are these others?
- How does the person's physical health affect him/her?
- How bad is the person's memory?

- Does the person's behavior or mood affect them or others?
- What kind of support does the person need? How many hours a day?
- How much time does the person's carer spend with them? Is this doable?
- How much care, overall, does the person need?

The example questions were adapted to the local context and routine data collection of each participating center.

Interviewers/raters were chosen by the leaders of each center according to which professionals usually see patients in their country; while in some instances these were general practitioners, in other settings psychiatric nurses or workers acted as interviewers or raters. Patients and their carers were previously unknown to interviewers/raters.

Interviews were conducted in the language of each center. The draft IDEAL schedule, its accompanying glossary, and a specially designed case report form were translated into the language of each center prior to data collection. Simple one-way translation techniques were considered sufficient in those centers where the study leads were part of the IDEAL steering committee (Spain, Netherlands, France and Germany), since project leaders in these centers were bilingual, since they had participated in the production of the project materials, and since the IDEAL schedule and accompanying glossary were produced to enable raters to use them in summarizing the results of their interview (i.e., not to be asked verbatim). Back-translation techniques were employed in the other field-sites.

Analyses

For focus groups, analyses were conducted based on the summary transcripts of the discussions to gain an overview of participants' responses; responses were ordered and grouped according to the different questions that were discussed. Since there was no theory being developed from the data, and data were relatively brief, more formal qualitative techniques were not appropriate.

Quantitative data analyses from the pilot- and field-studies were carried out in SPSS 17.0 (SPSS Inc., Chicago) (pilot-study) and SPSS 21.0 (IBM Corp, Armonk, NY) (field-study). Means and standard deviations were calculated for continuous demographic data of participants, as well as for mean scores on the IDEAL schedule, and counts and prevalence rates were calculated for categorical demographic data. To assess demographic differences between sites, ANOVA were performed for continuous variables, and chi-squared (χ^2) tests were employed for categorical variables.

ICCs (using a two-way random, absolute agreement, single measures model) were computed to assess inter-rater reliability of the IDEAL schedule [25]; ICCs were considered adequate if they were 0.64 or higher [26], assuming at least a medium level of correctness required for the classification. As secondary measures, Spearman's rho correlations were calculated between the total sum score of the IDEAL schedule and the CDR score, and Cronbach's alpha was computed to indicate internal consistency of the schedule.

RESULTS

Focus groups

Overall, the expert focus groups resulted in consensus that:

- There was a clear need for a comprehensive dementia staging scale, reflecting the individually required care profile.
- Staging is useful, in that it serves both the development of a 'common language', and the planning of health care services.
- Staging should be applicable to all types of dementias.
- Staging needs to be done with care and professionalism.
- An international standard is needed for staging instruments.

The focus groups also produced useful suggestions on the design and characteristics of an assessment tool, in particular that it should be (i) multi-axial (i.e., have multiple dimensions); (ii) developed in multiple languages; (iii) valid, reliable and feasible, in the hands of different professionals; (iv) simple and easy to use, not requiring extensive training prior to competent use (see [27] for support for the efficiency and efficacy of simple screening tests); (v) useable in clinical practice; and (vi) have goodness of fit with common dementia care practice.

To allow its use within countries worldwide and within a wide range of settings (whether clinical, non-clinical or research contexts), the schedule was developed to focus on the most important features of dementia, and to do so in a simple and easily understandable manner. To make justice to the clinical complexity of dementia, dimensions on the schedule

were developed to relate to a wide range of signs and symptoms, including behavioral, physical, functional, social, and care aspects, in addition to the cognitive symptoms more commonly included in previous dementia scales [28].

Face validity and content validity of the schedule were ensured through the consensus of a wide range of experts from several European countries about the schedule's format and content, i.e., the steering committee agreed on the schedule's format and content, based on the previous focus group discussions and vignette ratings that were conducted with a wide range of stakeholders involved in dementia care across Europe.

Pilot-study

ICCs for each of the draft IDEAL schedule's dimensions ranged between 0.16 and 0.73 during the first round of the pilot-study (see Table 1); between 0.44 and 0.78 during the second round (see Table 2); and between 0.33 and 0.90 during the third round (see Table 3). Several changes were made to the draft IDEAL schedule and its accompanying glossary over the three rounds of the pilot-study based on these inter-rater reliability findings (see Tables 1–3 and 7), in preparation for the larger field-study.

Field-study

Respondents

Tables 4 and 5 display the demographic characteristics of participants and their caregivers respectively during the field-study. Overall, across sites around two thirds of participants (65.1%) were female, and the

Table 1
Intra-class correlation coefficients[1] during first round of the pilot-study

IDEAL schedule dimensions	IDEAL members ($n = 5$)	IDEAL non-members ($n = 12$)
Activities of daily living	0.73 (0.50–0.91)	0.48 (0.27–0.77)
Physical health	0.73 (0.49–0.91)	0.58 (0.35–0.84)
Cognitive functioning	0.64 (0.35–0.87)	0.56 (0.34–0.82)
Behavioral disturbance	0.40 (0.14–0.73)	0.52 (0.31–0.80)
Social support/carer distress[2]	0.16 (−0.03–0.53)	0.23 (0.09–0.55)
Global rating of care need	0.52 (0.24–0.81)	0.46 (0.25–0.76)

[1]A two-way random, absolute agreement, single measures model was used; 95% confidence intervals are listed in brackets. [2]This dimension was divided into two dimensions following the first round of the pilot-study (see Table 2).

Table 2
Intra-class correlation coefficients[1] from second round of the pilot-study

IDEAL schedule dimensions	IDEAL members ($n = 5$)
Activities of daily living	0.74 (0.51–0.91)
Physical health	0.78 (0.57–0.93)
Cognitive functioning	0.72 (0.48–0.91)
Behavioral disturbance	0.58 (0.31–0.84)
Social support[2]	0.44 (0.18–0.76)
Carer distress	0.51 (0.24–0.81)
Global rating of care need[3]	0.73 (0.49–0.91)

[1]A two-way random, absolute agreement, single measures model was used; 95% confidence intervals are listed in brackets. [2]This dimension was divided into two sub-dimensions following the second round of the pilot-study (see Table 3). [3]This dimension was divided into three sub-dimensions following the second round of the pilot-study (see Table 3).

same was true for caregivers (64.1%). Most participants across sites displayed a mild to moderate degree of cognitive decline as measured by the CDR (76.0%). The majority of participants were living independently, either with family members or on their own (71.9%). Half of all caregivers who attended the interview with the patient were a child of the patient (49.8%), and a further third were the spouse or partner (32.7%).

However, there were some differences across field-sites in regards to patients' degree of cognitive decline ($p < 0.001$) (Romanian patients showed more severe cognitive decline than participants in other countries, and patients in Italy showed relatively milder cognitive decline), as well as in regards to patients' marital status across sites ($p = 0.04$), their living arrangements ($p < 0.001$), and their level of education ($p < 0.001$). The caregivers' gender and their relationship to the patient also varied across the field-sites ($p < 0.001$ for both variables).

There were missing data for the variables 'marital status', 'level of education', 'gender of carer', and 'carer's relationship to patient' (2 missing data points each), as well as for 'age' and 'CDR' (one missing data point each).

IDEAL scores

Table 6 shows the mean scores obtained for each of the IDEAL schedule's individual dimensions as well as the mean total scores in each of the nine field-sites. Results were generally comparable across sites, although ratings were relatively higher in Spain and Romania, and lower in South Korea. The 'Social support' dimension was rated lower across sites than the other dimensions.

Table 3
Intra-class correlation coefficients[1] from third round of the pilot-study

IDEAL schedule dimensions	Total (n = 15)	Spain (n = 7)	Netherlands (n = 6)	France (n = 2)
Activities of daily living	0.60 (0.39–0.84)	0.61 (0.36–0.85)	0.52 (0.26–0.81)	0.81 (0.43–0.95)
Physical health	0.86 (0.73–0.95)	0.84 (0.69–0.95)	0.87 (0.73–0.96)	0.90 (0.67–0.98)
Cognitive functioning	0.65 (0.45–0.87)	0.70 (0.48–0.90)	0.63 (0.37–0.86)	0.53 (−0.05–0.85)
Behavioral disturbance	0.46 (0.26–0.75)	0.52 (0.28–0.80)	0.41 (0.17–0.73)	0.33 (−0.27–0.77)
Social support[2] – Size of support network	0.53 (0.32–0.80)	0.63 (0.40–0.86)	0.63 (0.40–0.86)	0.73 (0.25–0.92)
Social support[2] – Social support	0.49 (0.28–0.77)	0.46 (0.22–0.76)	0.50 (0.24–0.79)	0.48 (−0.08–0.83)
Carer[3] – Time spent on care by carer	0.54 (0.33–0.81)	0.53 (0.28–0.81)	0.50 (0.25–0.80)	0.78 (0.34–0.94)
Carer[3] – Carer distress	0.64 (0.43–0.86)	0.58 (0.34–0.84)	0.64 (0.39–0.87)	0.83 (0.45–0.95)
Global rating of care need[4] – Amount of dementia-related care needed	0.62 (0.41–0.85)	0.66 (0.40–0.88)	0.63 (0.37–0.87)	0.68 (0.15–0.91)
Global rating of care need[4] – Number of hours needed for non-dementia care	0.67 (0.47–0.87)	0.64 (0.41–0.87)	0.69 (0.46–0.89)	0.59 (0.04–0.88)
Global rating of care need[4] – Type of care needed overall	0.61 (0.40–0.85)	0.69 (0.43–0.89)	0.51 (0.26–0.80)	0.66 (0.13–0.90)

[1]A two-way random, absolute agreement, single measures model was used; 95% confidence intervals are listed in brackets. [2]These two sub-dimensions were combined into one dimension following the third round of the pilot-study (see Table 6). [3]This dimension was rephrased as 'Non-professional care dimension' following the third round of the pilot-study (see Table 6). [4]This dimension was reconstructed into a 'Professional care dimension' following the third round of the pilot-study; the three sub-dimensions were also rephrased (see Table 7 and Supplementary Material).

Inter-rater reliability

ICCs for each of the IDEAL schedule's dimensions in each of the study centers ranged between 0.38 and 1.0 during the field-study, with 84.4% of ICCs over 0.7; ICCs for the total score (i.e., the sum score of all dimensions) ranged between 0.89 and 0.99 (see Table 7). Inter-rater reliability for all dimensions was considered to be adequate.

Inter-rater reliability results were generally similar across sites for most of the IDEAL schedule's dimension, with a few outliers (see Table 7). The largest differences across sites were for the 'Social support' dimension. This was in part a reflection of the fact that the anchor points and glossary wording of the 'Social support' dimension was revised following the field-study in Spain, Netherlands, France, and Germany to improve its inter-rater reliability, as results had not been satisfactory for this dimension; the inter-rater reliability results from Italy, Turkey, South Korea, Romania, and Serbia where the schedule was tested subsequently were very good (see Table 7).

There were also some differences between sites for the 'Behavioral disturbance' dimension. The dimension was therefore reworded as 'Behavioral and psychological symptoms' following the field-study, to clarify the dimension and to more accurately reflect its meaning.

Internal consistency

Internal consistency of the seven dimensions of the schedule was α = 0.80 across the nine field-sites.

Correlation analyses

There was a positive correlation between the total score of the IDEAL schedule and the CDR score ($\rho = 0.71$, $p < 0.001$) across the nine field-sites.

DISCUSSION

The IDEAL schedule was successfully developed, and then pilot-tested and field-tested in several countries worldwide. During its field-testing in clinical practice, the schedule met all requirements identified during the earlier focus groups; it therefore fulfills the preconditions to fill the gap that had been identified for a new staging schedule for dementia [17].

Strengths

The schedule is multi-dimensional (thereby relating to a wide range of symptoms and care aspects within dementia); it showed good content validity, high reliability and good feasibility during field-testing, and was found to be useful and easy to use by a range of professionals in several countries worldwide using different languages.

Table 4
Demographic characteristics of field-study participants (i.e., dementia patients)

	Total (n = 209)	Netherlands (n = 24)	Spain (n = 20)	France (n = 20)	Germany (n = 14)	Italy (n = 34)	Turkey (n = 32)	South Korea (n = 20)	Romania (n = 25)	Serbia (n = 20)
Gender										
Male	73 (34.9%)	11 (45.8%)	5 (25.0%)	7 (35.0%)	5 (35.7%)	14 (41.2%)	14 (43.8%)	4 (20.0%)	10 (40.0%)	3 (15.0%)
Female	136 (65.1%)	13 (54.2%)	15 (75.0%)	13 (65.0%)	9 (64.3%)	20 (58.8%)	18 (56.3%)	16 (80.0%)	15 (60.0%)	17 (20.0%)
Age, mean in years (SD)	77.5 (8.8)	80.9 (8.7)	78.3 (7.3)	76.9 (9.2)	82.3 (8.4)	77.7 (9.1)	73.9 (8.1)	75.5 (8.7)	77.9 (8.4)	76.7 (10.1)
Degree of cognitive decline										
Very mild to mild (CDR 0.5-1)	86 (41.1%)	12 (50.0%)	7 (35.0%)	6 (30.0%)	5 (35.7%)	23 (67.6%)	16 (50.0%)	8 (40.0%)	1 (4.0%)	8 (40.0%)
Moderate (CDR 2)	73 (34.9%)	6 (25.0%)	7 (35.0%)	7 (35.0%)	7 (50.0%)	7 (20.6%)	12 (37.5%)	11 (55.0%)	10 (40.0%)	6 (30.0%)
Severe (CDR 3 or above)	49 (23.4%)	6 (25.0%)	6 (30.0%)	7 (35.0%)	1 (7.1%)	4 (11.8%)	4 (12.5%)	1 (5.0%)	14 (56.0%)	6 (30.0%)
Marital status										
Married	117 (56.0%)	15 (62.5%)	11 (55.0%)	12 (60.0%)	7 (50.0%)	23 (67.6%)	20 (62.5%)	7 (35.0%)	17 (68.0%)	5 (25.0%)
Co-habiting	1 (0.5%)	0	0	0	0	0	0	1 (5.0%)	0	0
Divorced	10 (4.8%)	1 (4.2%)	2 (10.0%)	4 (20.0%)	0	1 (2.9%)	1 (3.1%)	0	0	1 (5.0%)
Widowed/partner deceased	76 (36.4%)	8 (33.3%)	7 (35.0%)	4 (20.0%)	6 (42.9%)	10 (29.4%)	10 (31.3%)	10 (50.0%)	8 (32.0%)	13 (65.0%)
Single/unmarried	3 (1.4%)	0	0	0	1 (7.1%)	0	1 (3.1%)	0	0	1 (5.0%)
Living arrangements										
Independent, alone, no day care	25 (12.0%)	6 (25.0%)	0	6 (30.0%)	1 (7.1%)	4 (11.8%)	4 (12.5%)	4 (20.0%)	0	0
Independent, alone, with day care	15 (7.2%)	0	0	8 (40.0%)	2 (14.3%)	1 (2.9%)	3 (9.4%)	0	1 (4.0%)	0
Independent, with others, no day care	67 (32.1%)	10 (41.7%)	2 (10.0%)	4 (20.0%)	2 (14.3%)	28 (82.4%)	9 (28.1%)	11 (55.0%)	10 (40.0%)	1 (5.0%)
Independent, with others, with day care	43 (20.6%)	1 (4.2%)	4 (20.0%)	2 (10.0%)	8 (57.1%)	1 (2.9%)	10 (31.3%)	3 (15.0%)	10 (40.0%)	4 (20.0%)
Care/residential home	21 (10.0%)	2 (8.3%)	6 (30.0%)	0	0	0	6 (18.8%)	2 (10.0%)	0	5 (25.0%)
Nursing home	23 (11.0%)	5 (20.8%)	8 (40.0%)	0	1 (7.1%)	0	0	0	0	9 (45.0%)
Level of education										
Primary or special education school	129 (61.7%)	10 (41.7%)	16 (80.0%)	13 (65.0%)	4 (28.6%)	28 (82.4%)	21 (65.6%)	14 (70.0%)	15 (60.0%)	8 (40.0%)
Secondary or vocational school	56 (26.8%)	12 (50.0%)	3 (15.0%)	4 (8.0%)	5 (35.7%)	6 (17.6%)	4 (12.5%)	6 (30.0%)	5 (20.0%)	11 (55.0%)
University/tertiary education	11 (5.3%)	1 (4.2%)	1 (5.0%)	3 (15.0%)	4 (28.6%)	0	2 (6.3%)	0	0	0
Other	11 (5.3%)	0	0	0	0	0	5 (15.6%)	0	5 (20.0%)	1 (0.1%)

Numbers do not always add up to total number of participants due to missing data.

Table 5
Demographic characteristics of dementia patients' caregivers during the field-study

	Total (n = 217)	Netherlands (n = 32)	Spain (n = 20)	France (n = 20)	Germany (n = 14)	Italy (n = 34)	Turkey (n = 32)	South Korea (n = 20)	Romania (n = 25)	Serbia (n = 20)
Gender										
Male	76 (35.0%)	15 (46.9%)	6 (30.0%)	9 (45.0%)	5 (35.7%)	9 (26.5%)	12 (37.5%)	7 (35.0%)	9 (36.0%)	4 (20.0%)
Female	139 (64.1%)	17 (53.1%)	14 (70.0%)	10 (50.0%)	9 (64.3%)	25 (73.5%)	20 (62.5%)	12 (60.0%)	16 (64.0%)	16 (80.0%)
Relationship to patient										
Spouse or partner	71 (32.7%)	12 (37.5%)	9 (45.0%)	10 (50.0%)	2 (14.3%)	14 (41.2%)	12 (37.5%)	8 (40.0%)	1 (4.0%)	3 (15.0%)
Child of patient	108 (49.8%)	19 (59.4%)	9 (45.0%)	5 (25.0%)	11 (78.6%)	15 (44.1%)	19 (59.4%)	10 (50.0%)	14 (56.0%)	6 (30.0%)
Other family member	15 (6.9%)	1 (3.1%)	0	4 (20.0%)	0	1 (2.9%)	1 (3.1%)	0	7 (28.0%)	1 (5.0%)
Other (e.g., friend)	21 (9.7%)	0	2 (10.0%)	0	1 (7.1%)	4 (11.8%)	0	1 (5.0%)	3 (12.0%)	10 (50.0%)

Numbers do not always add up to total number of participants due to missing data, and in some cases more than one carer attending the interview (in the Netherlands).

Table 6
Mean scores obtained for the IDEAL schedule in field-sites

IDEAL schedule dimensions	Total (n=209)	Spain (n=20)	Netherlands (n=24)	France (n=20)	Germany (n=14)	Italy (n=34)	Turkey (n=32)	South Korea (n=20)	Romania (n=25)	Serbia (n=20)
Activities of daily living	2.78 (SD=1.59)	3.0 (SD=1.62)	2.96 (SD=1.49)	2.85 (SD=1.57)	2.93 (SD=1.21)	2.47 (SD=1.93)	2.56 (SD=1.37)	1.90 (SD=1.41)	3.44 (SD=1.58)	3.05 (SD=1.50)
Physical health	2.22 (SD=1.40)	2.30 (SD=1.42)	1.67 (SD=1.09)	2.25 (SD=1.48)	1.71 (SD=1.38)	3.06 (SD=1.61)	2.44 (SD=1.29)	1.55 (SD=0.69)	2.36 (SD=1.55)	1.90 (SD=1.12)
Cognitive functioning	2.87 (SD=1.27)	2.95 (SD=1.50)	3.04 (SD=1.30)	3.05 (SD=1.23)	2.71 (SD=0.83)	2.82 (SD=1.38)	2.5 (SD=1.22)	2.20 (SD=0.77)	3.56 (SD=1.23)	2.95 (SD=1.32)
Behavioral and psychological symptoms[1]	2.17 (SD=1.27)	2.10 (SD=1.45)	2.08 (SD=0.97)	2.15 (SD=1.09)	2.14 (SD=1.23)	1.91 (SD=1.42)	2.22 (SD=1.07)	1.80 (SD=0.89)	3.24 (SD=1.42)	1.75 (SD=1.16)
Social support[2]	1.13 (SD=1.09)	1.15 (SD=0.93)	1.08 (SD=0.83)	1.35 (SD=0.99)	1.14 (SD=1.10)	0.44 (SD=1.11)	1.31 (SD=1.28)	1.60 (SD=0.75)	1.36 (SD=1.0)	1.10 (SD=1.29)
Non-professional care	2.52 (SD=1.47)	3.50 (SD=1.15)	2.33 (SD=1.03)	2.10 (SD=1.36)	2.86 (SD=1.29)	2.59 (SD=1.77)	2.36 (SD=1.36)	1.03 (SD=1.24)	3.22 (SD=1.40)	2.68 (SD=1.09)
Professional care	1.81 (SD=1.45)	3.17 (SD=1.58)	2.68 (SD=1.62)	1.53 (SD=1.16)	2.21 (SD=1.26)	0.83 (SD=1.03)	1.57 (SD=0.86)	0.90 (SD=1.41)	2.20 (SD=1.11)	1.90 (SD=1.49)
Total score	15.46 (SD=6.40)	18.17 (SD=5.24)	15.85 (SD=6.25)	15.18 (SD=6.32)	15.71 (SD=5.02)	14.13 (SD=7.25)	14.72 (SD=6.15)	10.98 (SD=5.0)	19.38 (SD=6.54)	15.33 (SD=5.59)

[1]This dimension was reworded from 'Behavioral disturbance' to 'Behavioral and psychological symptoms' following the field-study. [2]The anchor points and glossary wording of this dimension were rephrased following the field-study in Spain, France, Netherlands and Germany.

Table 7
Intra-class correlation coefficients[1] during field-study

IDEAL schedule dimensions	Total (n=209) ICCs[1] (95% CI)	Spain (n=20) ICCs[1] (95% CI)	Netherlands (n=24) ICCs[1] (95% CI)	France (n=20) ICCs[1] (95% CI)	Germany (n=14) ICCs[1] (95% CI)	Italy (n=34) ICCs[1] (95% CI)	Turkey (n=32) ICCs[1] (95% CI)	South Korea (n=20) ICCs[1] (95% CI)	Romania (n=25) ICCs[1] (95% CI)	Serbia (n=20) ICCs[1] (95% CI)
Activities of daily living	0.92 (0.89–0.94)	0.89 (0.75–0.96)	0.91 (0.81–0.96)	0.93 (0.82–0.97)	0.56 (0.06–0.83)	0.98 (0.96–0.99)	0.87 (0.76–0.94)	0.87 (0.69–0.95)	0.96 (0.89–0.98)	0.98 (0.94–0.99)
Physical health	0.88 (0.85–0.91)	0.79 (0.55–0.91)	0.82 (0.63–0.92)	0.86 (0.68–0.94)	0.65 (0.23–0.87)	0.96 (0.93–0.98)	0.81 (0.66–0.91)	0.49 (0.08–0.76)	0.99 (0.98–1.0)	0.98 (0.95–0.99)
Cognitive functioning	0.93 (0.91–0.94)	0.95 (0.88–0.98)	0.92 (0.84–0.97)	0.90 (0.77–0.96)	0.55 (0.08–0.82)	0.98 (0.97–0.99)	0.85 (0.72–0.93)	0.92 (0.80–0.97)	0.97 (0.94–0.99)	0.97 (0.93–0.99)
Behavioral and psychological symptoms[2]	0.84 (0.80–0.88)	0.83 (0.63–0.93)	0.62 (0.30–0.82)	0.74 (0.45–0.89)	0.86 (0.63–0.95)	0.99 (0.97–0.99)	0.77 (0.58–0.88)	0.66 (0.31–0.85)	0.92 (0.83–0.96)	0.76 (0.47–0.90)
Social support[3]	0.78 (0.72–0.83)	0.59 (0.20–0.81)	0.68 (0.39–0.85)	0.38 (−0.09–0.70)	0.73 (0.34–0.90)	0.99 (0.97–0.99)	0.71 (0.49–0.85)	0.81 (0.52–0.92)	1.0 (1.0–1.0)	0.82 (0.60–0.92)
Non-professional care	0.90 (0.87–0.92)	0.96 (0.89–0.98)	0.86 (0.69–0.94)	0.54 (0.16–0.79)	0.70 (0.32–0.89)	1.0 (0.99–1.0)	0.82 (0.66–0.91)	0.96 (0.91–0.98)	0.99 (0.97–0.99)	0.93 (0.75–0.97)
Professional care	0.98 (0.97–0.98)	0.98 (0.95–0.99)	0.99 (0.97–0.99)	0.91 (0.79–0.96)	0.91 (0.75–0.97)	1.0 (0.99–1.0)	0.90 (0.81–0.95)	0.99 (0.97–1.0)	1.0 (1.0–1.0)	0.96 (0.90–0.99)
Total score	0.96 (0.94–0.97)	0.90 (0.76–0.96)	0.92 (0.84–0.97)	0.89 (0.74–0.96)	0.93 (0.64–0.98)	0.99 (0.99–1.0)	0.96 (0.92–0.98)	0.92 (0.81–0.97)	0.99 (0.97–1.0)	0.94 (0.79–0.98)

[1]Intra-class correlation coefficient; a two-way random, absolute agreement, single measures model was used. [2]This dimension was reworded from 'Behavioral disturbance' to 'Behavioral and psychological symptoms' following the field-study. [3]The anchor points and glossary wording of this dimension were rephrased following the field-study in Spain, France, Netherlands and Germany. CI, confidence interval.

Inter-rater reliability of the seven dimensions of the schedule were adequate to excellent in each of the nine countries in which the schedule was field-tested, and inter-rater reliability for the total sum score was excellent in each of the study centers. Face validity and content validity of the schedule were established by the expert views of the multi-disciplinary steering committee and their colleagues that the schedule adequately addressed relevant dementia concepts, and that the content of the schedule was comprehensive.

Secondary analyses showed that the total score of the IDEAL schedule correlated positively with the CDR score, providing preliminary evidence for concurrent validity of the schedule. The schedule also showed good internal consistency. However, very high internal consistency should not be expected for the schedule, as its different dimensions may not necessarily relate to each other positively; for example, activities of daily living may be unrelated to physical ill-health, or the professional and non-professional care dimensions may be inversely related.

Limitations and future work

Limitations of the study include that participants were selected into the field-study using non-randomization methods, and that simple translation techniques were employed in some of the field-sites. However, since within psychometric testing the focus is more upon substantive responses than upon the representativeness of participants, and since the schedule was not administered verbatim during field-study interviews, this should have had a minimal effect on the results. Further limitations are that inter-rater reliability was assessed using a silent observer (who may be affected by the interviewer's manner of conducting the interview), and that no quantitative measurement (such as Lawshe or Rand Delphi) was used to assess content validity.

Hence, although this study is a first step toward showing reliability and validity of the schedule, further work is needed to test the psychometric properties of the schedule, such as test-retest reliability, concurrent validity, predictive validity, discriminant validity, or sensitivity to change. Another line of work for the future may be to assess the possibility of introducing scores for a patient domain and a care domain (in addition to scores for individual dimensions and a total sum score), for example through factor analyses. While the IDEAL group itself is planning further psychometric testing of the schedule (and indeed is already doing so in a wider range of countries, including further

countries outside of Europe), we also encourage other research groups to do so using the schedule in our Supplementary Material, to provide a larger evidence base. For this, it will be important to also test the schedule in other regions of the world, to ensure that the schedule is appropriate for use worldwide, thereby fulfilling its aims of being a global staging scale [17].

Importantly, to fully address the issue of staging within dementia care, work is currently being carried out by the IDEAL group to produce a menu of procedures or interventions corresponding to each of the different stages of dementia (according to the IDEAL schedule); this should further improve and add to the schedule. Eventually, a menu of interventions will therefore be attached to the schedule, with a range of interventions being suggested for different dementia symptom and severity patterns. For example, different interventions would be proposed for patients who primarily display severe cognitive deficiencies compared to those who display mainly behavioral problems (as measured by the IDEAL schedule), or compared to those patients who display high scores on all dimensions of the IDEAL schedule. Practitioners would then be able to choose from this standard of interventions, depending on the patient's symptomatology and severity levels, as well as their setting and the resources they have available.

Implications

We hope that ultimately the IDEAL schedule may be able to facilitate the clinical and (non-clinical) social management of care for people with dementia to be organized more uniformly, efficiently, and effectively. It may do so in several ways. First, use of the schedule might give professional caregivers the opportunity to have a better understanding of patients' condition across multiple areas, enabling them to focus on the most important problems and to act proactively. Second, use of the schedule may serve as a 'common language' between caregivers and professionals. Third, at a global level, it may be used to increase insight into the amount of possible mismatches between the required and available resources; this might make it easier to allocate monetary and manpower resources to the people who need it most. Fourth, it may be used to address caregiver burden and to help adjust the planning of care according to the wishes and needs of the carers.

Whether the schedule is able to fulfill these roles remains to be proven. However, if it is able to do so, it will be an important tool in helping to improve the care for people with dementia cross-nationally, and in

helping to reduce some of the large burden dementia inflicts on the patients themselves, as well as on their families, caregivers and society at large.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The supplementary material is available in the electronic version of this article: http://dx.doi.org/10.3233/JAD-141599.

REFERENCES

[1] Alzheimer's Disease International (ADI), World Health Organization (WHO) (2012) *Dementia: A public health priority*, WHO, Geneva.

[2] Prince MJ, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP (2013) The global prevalence of dementia: A systematic review and meta-analyses. *Alzheimers Dement* **9**, 63-75.

[3] Fortinsky RH, Downs M (2014) Optimizing person-centered transitions in the dementia journey: A comparison of national dementia strategies. *Health Aff* **33**, 4566-4573.

[4] Edge SB, Compton CC (2010) The American Joint Committee on Cancer: The 7th edition of the AJCC Cancer Staging Manual and the future of TNM. *Ann Surg Oncol* **17**, 1471-1474.

[5] Cosci F, Fava GA (2013) Staging of mental disorders: Systematic review. *Psychother Psychosom* **82**, 20-34.

[6] Kapczinski F, Dias VV, Kauer-Sant'Anna M, Frey BN, Grassi-Oliveira R, Colom F, Berk M (2009) Clinical implications of a staging model for bipolar disorders. *Expert Rev Neurother* **9**, 957-966.

[7] McGorry PD, Hickie IB, Yung AR, Pantelis C, Jackson HJ (2006) Clinical staging of psychiatric disorders: A heuristic framework for choosing earlier, safer and more effective interventions. *Aust NZ J Psychiatry* **40**, 616-622.

[8] McNamara RK, Nandagopal JJ, Strakowski SM, DelBello MP (2010) Preventive strategies for early-onset bipolar disorder: Towards a clinical staging model. *CNS Drugs* **24**, 983-996.

[9] Vieta E, Reinares M, Rosa AR (2010) Prognosis and staging in bipolar disorder. *Actas Esp Psiquiatr* **38**(Suppl 3), 35-38.

[10] Agius M, Goh C, Ulhaq S, McGorry P (2010) The staging model in schizophrenia, and its clinical implications. *Psychiatr Danub* **22**, 211-220.

[11] Fava GA, Kellner R (1993) Staging: A neglected dimension in psychiatric classification. *Acta Psychiatr Scand* **87**, 225-230.

[12] Lieberman JA, Perkins D, Belger A, Chakos M, Jarskog F, Boteva K, Gilmore J (2001) The early stages of schizophrenia: Speculations on pathogenesis, pathophysiology, and therapeutic approaches. *Biol Psychiatry* **50**, 884-897.

[13] Fava GA, Tossani E (2007) Prodromal stage of major depression. *Early Interv Psychiatry* **1**, 9-18.

[14] Hetrick SE, Parker AG, Hickie IB, Purcell R, Yung AR, McGorry PD (2008) Early identification and intervention in depressive disorders: Towards a clinical staging model. *Psychother Psychosom* **77**, 263-270.

[15] Balsis S, Miller TM, Benge JF, Doody RS (2011) Dementia staging across three different methods. *Dement Geriatr Cogn Disord* **31**, 328-333.

[16] Piguet O, Hornberger M, Mioshi E, Hodges JR (2011) Behavioral-variant frontotemporal dementia: Diagnosis, clinical staging, and management. *Lancet Neurol* **10**, 162-172.

[17] Olde Rikkert MGM, Tona KD, Janssen L, Burns A, Lobo A, Robert P, Sartorius N, Stoppe G, Waldemar G (2011) Validity, reliability and feasibility of clinical staging scales in dementia: A systematic review. *Am J Alzheimers Dis Other Demen* **26**, 357-365.

[18] Byrne EJ, Benoit M, Lopez Arrieta JM, Geraldi C, Koopmans R, Rolland Y, Sartorius N, Stoppe G, Robert P (2008) For whom and for what the definition of severe dementia is useful: An EDCON consensus. *J Nutr Health Aging* **12**, 714-719.

[19] Olde Rikkert MG, van der Vorm A, Burns A, Dekkers W, Robert P, Sartorius N, Selmes J, Stoppe G, Vernooij-Dassen M, Waldemar G (2008) Consensus statement on genetic research in dementia. *Am J Alzheimers Dis Other Demen* **23**, 262-626.

[20] Stoppe G, on behalf of the European Dementia Consensus Network (2008) *Competence Assessment in Dementia*. Springer, Wien/New York.

[21] Waldemar G, Phung KT, Burns A, Georges J, Hansen FR, Iliffe S, Marking C, Rikkert MO, Selmes J, Stoppe G, Sartorius N (2007) Access to diagnostic evaluation and treatment for dementia in Europe. *Int J Geriatr Psychiatry* **22**, 47-54.

[22] Perneczky R, Wagenpfeil S, Komossa K, Grimmer T, Diehl J, Kurz A (2006) Mapping scores onto stages: Mini-mental state examination and clinical dementia rating. *Am J Geriatr Psychiatry* **14**, 139-144.

[23] Rabheru K (2007) Disease staging and milestones. *Can J Neurol Sci* **34**, S62-S66.

[24] Langbart C (2002) Diagnosing and treating Alzheimer's disease: A practitioner's overview. *J Am Acad Nurse Pract* **14**, 103-109.

[25] Shrout PE, Fleiss JL (1979) Intraclass correlations: Uses in assessing rater reliability. *Psychol Bull* **2**, 420-428.

[26] Ellis JL (2013) Probability interpretations of intraclass reliabilities. *Stat Med* **32**, 4596-4608.

[27] Ventura T, De-la-Cámara C, Lopez-Anton R, Santabárbara J, Marcos G, Quintanilla MA, Campayo A, Lobo E, Roy JF, Saz P, Olaya B, Haro JM, Rodríguez-Mañas L, Sartorius N, Lobo A (2013) Usefulness of 2 questions about age and year of birth in the case-finding of dementia. *J Am Med Dir Assoc* **14**, 627.e7-12.

[28] Saz P, López-Antón R, Dewey ME, Ventura T, Martín A, Marcos G, De La Cámara C, Qunitanilla MA, Quetglas B, Bel M, Barrera A, Lobo A (2009) Prevalence and implications of psychopathological non-cognitive symptoms in dementia. *Acta Psychiatr Scand* **119**, 107-116.