Dauphinot *et al. Alzheimer's Research & Therapy* (2024) 16:198
https://doi.org/10.1186/s13195-024-01567-9

Alzheimer's Research & Therapy

RESEARCH Open Access

Reliability of the assessment of the clinical dementia rating scale from the analysis of medical records in comparison with the reference method

Virginie Dauphinot[1,3*], Sylvain Calvi[1], Claire Moutet[1], Jing Xie[1], Sophie Dautricourt[1,2], Anthony Batsavanis[1], Pierre Krolak-Salmon[1†] and Antoine Garnier-Crussard[1,2†]

Abstract

Background The Clinical Dementia Rating (CDR) scale allows to detect the presence of dementia and to assess its severity, however its evaluation requires a significant time (45 min). We evaluated the agreement between two methods of collection of the CDR: face-to-face interview or based on the information available in the patient's medical record.

Methods The CLIMER study was conducted among patients attending a memory center. The CDR scale was evaluated during face-to-face interviews between neuropsychologists and patients and their caregivers and based on blind analysis of the information of the patients' medical record by neuropsychologists. The agreement of the CDR sum of boxes (CDR-SB), the 5-point scale CDR and the different domains of the CDR evaluated between the different methods was measured using intraclass correlation (ICC) coefficient, Bland and Altman method, and linearly weighted Kappa.

Results The study included 139 patients (means ± SD age 80.1 ± 6, 58.3% women, 71.9% with dementia). The ICC for the CDR-SB score assessed by face-to-face and with all the information available in the patient's medical record was 0.95 (95% CI: 0.93–0.97). The mean difference between the CDR-SB score assessed by face-to-face and with the medical record was 0.098 ± 1.036, and 92.4% of the patients lay within the 95% limits of agreement. The ICC for the 5-point scale CDR assessed by face-to-face and with the patient's medical record was 0.92 (95% CI: 0.88–0.95) when all the available information of the patient's medical record was used. The linear weighted Kappa coefficients was 0.79 (95% CI: 0.68–0.91) for the 5-point scale CDR comparison between the two evaluation methods. The analysis by domain of the CDR showed ICC ranging from 0.65 to 0.91 depending of the domains and the methods of evaluation.

†Pierre Krolak-Salmon and Antoine Garnier-Crussard are senior authors.

*Correspondence:
Virginie Dauphinot
virginie.dauphinot@chu-lyon.fr

Full list of author information is available at the end of the article

Conclusion This study showed an excellent level of agreement of the evaluation of the CDR- SB and the 5-point scale CDR when using all the information of the patient's medical record compared to the face-to-face interview.

Trial registration https//clinicaltrials.gov/ct2/show/NCT04763941 Registration Date 02/17/2021.

Keywords Clinical dementia rating, Alzheimer's disease, Reliability, Neuropsychology, Neurocognitive disorders

Background

Assessing the degree of severity of cognitive impairment is essential in the follow-up of patients with neurocognitive disorders (NCD) and to assess the effectiveness of therapeutics. The Clinical Dementia Rating (CDR) scale, developed by Hughes et al. in 1982, allows to clinically detect the presence of dementia such as Alzheimer's disease (AD) and to assess its severity [1]. Among the 2 scores that can be calculated from the CDR scale (the 5-point scale CDR and the CDR-Sum of Boxes (CDR-SB)), the CDR-SB score offers greater precision to determinate the stages of dementia and monitoring their evolution [2].

The CDR allows to measure not only cognition, but also functional autonomy, which makes it more comprehensive than a purely cognitive assessment. The CDR is now widely accepted in clinical setting and medical research as a valid and reliable evaluation measure [3], and it is particularly useful for studies at early stages of AD [4]. For example, change in the CDR-SB was the main outcome in the recent ENGAGE/EMERGE (Aducanumab) [5] and CLARITY AD (Lecanemab) [6] studies, and a secondary outcome of TRAILBLAZER-ALZ 2 study (Donanemab) [7]. However, in routine care, the systematic collection of the CDR scale presents certain limits due to the time required for its realization (approximately 45 min), and because the questions are sometimes redundant with other evaluations. Therefore, the question arises to determine whether there is another method to more systematically assess the CDR.

Previous studies have shown significant correlations between CDR scores and a selection of neuropsychological tests (number of words learned, eight MMSE orientation items, Verbal Fluency, CERAD Boston Naming test and non-orientation items of MMSE) [8, 9]. In particular, Fillenbaum et al. highlighted correlations between neuropsychological tests evaluating specific domains (memory, orientation, problem solving) with the CDR-SB score evaluating these same domains [8]. In addition, Fillenbaum et al. indicate that memory underpins daily life so much that when it becomes impaired, other aspects of cognition are also impaired and difficulties can be observed in outdoor activities, social and leisure activities, as well as in the performance of household tasks. In Perneczky et al., the authors determined to what extent thresholds applied to the Mini-Mental State examination (MMSE) could match the stages of CDR, in order to save time and to detect dementia without the need for an informant [9]. They showed that the classification of patients according to their MMSE allowed to determine the stages of mild, moderate and severe major NCD (or dementia) defined according to the CDR scale, nevertheless the correspondence between the classification of the MMSE and the CDR in earlier stage was low.

Thus, as studies have shown good correspondences between the CDR scale and the scales measuring cognitive and neuropsychological performance, and that several functional and neuropsychological scales are often evaluated in the context of memory consultations, we aimed to study whether the evaluation of the CDR scale based on the information already available in the patient's medical record would provide a reliable measure compared to the evaluation of the CDR scale by face-to-face interview in consultation. In particularly, we studied whether the CDR-SB scores obtained with these two ways would have a good degree of agreement.

Methods

Study design and setting

The CLIMER (Clinical Dementia Rating Medical Record) study is an observational cross-sectional study. The study focused on data collected during memory consultations and as part of the MEMORA cohort [10]. The study has been carried out between 2019 and 2022 at the Memory Consultation of the Clinical and Research Memory Center of Lyon, at the Charpennes Hospital (Hospices Civils de Lyon), France. The study has been performed using the medical record of patients attending the memory center; the information were collected during routine care and from the MEMORA cohort.

Population study and sample size

The population study included consecutive patients attending a memory center, accompanied by their caregiver, whatever the neurocognitive etiology.

The patient's medical records were selected consecutively over the study period, and according to the following criteria: patient of the memory center having a face-to-face interview allowing the assessment of the CDR-SB; patient with subjective cognitive decline (SCD) [11] or a neurocognitive disorders (mild cognitive impairment (MCI) or dementia) [12, 13] with an MMSE ≥ 12/30; and patients included in the MEMORA cohort. Patients or caregivers who did not want their data to be used for

research purposes as part of the MEMORA cohort were not included. This research was an observational-type study (without intervention and new assessment requiring a new visit of the patient and his/her caregiver) which is part of the pathway of patients attending a memory center.

The number of subjects needed was 144 patients, estimated on the basis of an expected Cohen's Kappa of 0.8, a standard deviation of 0.05, a proportion of patients with dementia expected at 40% during the face-to-face assessment and a similar proportion of dementia (40%) expected during the evaluation on the patient's medical record, and an absolute precision of 0.1 (sskdlg function using Stata).

Conduct of the evaluation

The design of the study is presented in Fig. 1.

The first CDR assessment was carried out during the face-to-face interview between a trained neuropsychologist, the patient and his/her caregiver, as part of the routine care. The second CDR assessment was carried out on the patient's medical record based on the information available and for which the date does not exceed a period of 2 months with the date of the first CDR assessment. This maximum period of 2 months was set to limit the possibility of changes in the patient's state of health, which would make the comparison irrelevant.

The CDR assessment on patient's medical record was performed by neuropsychologists who read and analyzed the information on the medical records based on what they deemed necessary to determine the CDR-SB scores. The neuropsychologists who assessed the CDR scale have all received a 9-hour training course aimed at standardizing practices and therefore inter-rater reliability (https://knightadrc.wustl.edu/cdr/cdr.htm).

This new assessment was performed blind to the first CDR, and the different CDR scores were also assessed by a different neuropsychologist; procedures were implemented so that the neuropsychologists in charge of analyzing the patient's medical record did not have access to the first assessment of the CDR.

The assessment of the CDR on patient's medical record was based on an implicit approach (based on judgment) from the information available rather than an explicit approach (based on strict criteria). The implicit approach took into account that different scales may be used to conduct the patient assessment but that they provide similar information. The implicit approach was therefore closer to the reality as the assessments, available in medical records, may vary from a patient to another. In addition, as part of the standard CDR interview assessment, the scoring appeals to the judgment of the neuropsychologist during the semi-structured interview. Nevertheless, all the information used to determine the CDR scores were collected in a specific table in order to identify whether similar scales were used systematically to evaluate the patients.

In the memory center where the study was conducted, data relating to the evaluation of neuropsychological and functional performances are collected during the memory consultation and during interviews between physicians, nurses, neuropsychologists, the patients and their

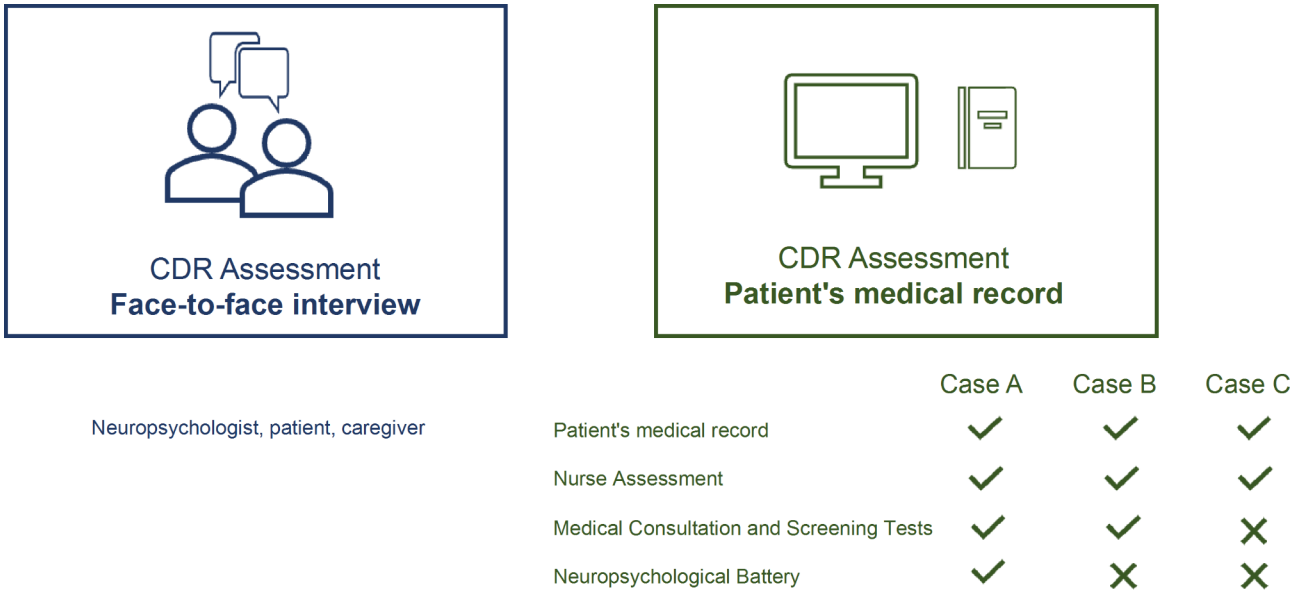

Fig. 1 Design of the study. In both cases, assessment was done by a trained blinded neuropsychologist. For the Clinical Dementia Rating (CDR) assessment based on medical record, nurse assessment contains information on global cognition (e.g. Mini Mental Stage Examination) and autonomy (e.g. Instrumental Activities of Daily Living), medical consultation contains cognitive screening tests (e.g. Frontal Assessment Battery, 5 words…) and comprehensive neuropsychological battery contains specific tests in cognitive subdomains

primary caregivers. The data are entered in an Electronic Health Record system by nursing staff, paramedics and consultation secretaries.

Study outcomes and patient's characteristics

The primary outcome was the CDR-SB score (score out of 18) obtained (1) during a face-to-face interview at the memory consultation, and (2) obtained with all the information available in the patient's medical record for the same patients (case A). This information could be results of the nurse assessment (global cognition and autonomy scales), of the medical consultation (cognitive screening tests…) as well as comprehensive neuropsychological battery.

The CDR-SB score allows to distinguish the following classes: 0 (normal cognition), 0.5-4 (questionable cognitive impairment), 0.5–2.5 (questionable impairment), 3–4 (very mild major NCD/or very mild dementia), 4.5- 9 (mild major NCD/or mild dementia), 9.5–15.5 (moderate major NCD/or moderate dementia), 16–18 (severe major NCD/or severe dementia) [14].

In the CDR scale, the patient's abilities are assessed based on the information available in six different areas: three areas reflect cognitive abilities: memory, orientation and judgment skills and problem solving, the other three areas reflect the actions of everyday life: outdoor and social activities, domestic and leisure activities, as well as personal care.

The secondary outcomes included:

- The CDR-SB score evaluated on file using all the information available in the patient file, except the comprehensive neuropsychological assessment (case B);
- The CDR-SB score assessed on file using all the information available in the patient file, except the comprehensive neuropsychological assessment, and the medical doctor's clinical examination, including cognitive screening tests (case C) ;
- The 5-point scale CDR with the five classes: 0: normal, 0.5: very mild dementia, 1: mild dementia, 2: moderate dementia, and 3: severe dementia.
- The CDR sub-scores for each of the six cognitive and functional areas evaluated in the CDR carried out in face-to-face interview and on patient's medical record (memory, guidance and judgment, participation in community life, home and leisure occupations and personal care) were considered.

The possible correspondence between the domains of the CDR and the data used to evaluate the CDR with the patient's medical record was described before the evaluation (supplement Table 1), nevertheless the neuropsychologists remained free to use the information available and that were relevant for them.

In addition to the information used from the patient's medical record to assess the CDR-SB scores and CDR sub-scores, the others data collected from the MEMORA cohort were: sociodemographic data (age, sex, education), the current living situation of the patients, the diagnostic stage and probable etiology if determined by the physician in charge of the patient, the global cognitive assessment (MMSE, score/30 [15]), the functional assessment (IADL, score/8 [16], DAD-6, score /18 [17], a scale assessing the caregiver burden (mini-Zarit burden interview, score/7 [18]), and the assessment of behavioral and psychological symptoms (score from the Neuropsychiatry Inventory (NPI, score/144 [19]).

Statistical analyses

The characteristics of the study population were described globally.

The degree of agreement between the CDR-SB scores assessed by face-to-face interview or with the patient's medical record according to the different cases A, B and C was measured using: the intra-class coefficient (ICC) (("two-way mixed effects, absolute agreement, and multiple raters/measurement" form). The level of agreement based on the ICC was considered poor with ICC was <0.5, moderate when ICC was between 0.5 and 0.75, good when ICC was between 0.75 and 0.9, and excellent when ICC was ≥ 0.9 ; and the Bland and Altman method including the Bland and Altman plots for each comparison [20].

The degree of agreement between the face-to-face interview and the patient's medical records to evaluate the 5-point scale CDR and the CDR sub-scores (for each of the 6 areas) was measured using: the intra-class coefficient (ICC) (("two-way mixed effects, absolute agreement, multiple raters/measurement" form), and the linearly weighted Cohen's Kappa coefficient [21]. The degree of agreement using the Kappa coefficients was interpreted as for a coefficient <0: disagreement, 0.1–0.2: very weak agreement, 0.21–0.4: weak agreement, 0.41–0.6: moderate agreement, 0.61–0.80: strong agreement, ≥ 0.8-1: almost perfect agreement.

The information used by the neuropsychologists in the patient's medical records was also described by domains.

To evaluate the feasibility of evaluating CDR using the patient's medical record, the duration of evaluation of the CDR (in minutes) was compared between the methods of assessment using paired t-test. The proportions of evaluation of the CDR scale using the patient's medical records were described for each case (A, B and C), as the information was not always available to evaluate the CDR in each case.

Table 1 Characteristics of the patients' sample

n = 139	*n* (%) or mean ± SD
Age (years)	80.14 ± 6.20
Gender	
Woman	81 (58.3)
Man	58 (41.7)
Education	
Primary	48 (34.5)
Secondary	61 (43.9)
Tertiary	30 (21.6)
Current living situation	
At home	127 (91.4)
Others	12 (8.6)
Clinical diagnosis stage	
Dementia	100 (71.9)
Mild Cognitive Impairment	33 (23.7)
Subjective Cognitive Decline	6 (4.3)
Etiology	
Alzheimer's disease*	121 (87)
Others†	18 (12.9)
MMSE (score/30)	20.68 ± 4.45
IADL (score/8) (*n* = 133)	4.20 ± 2.06
DAD6 (score/18) (*n* = 124)	7.49 ± 5.45
NPI (score/144) (*n* = 126)	17.27 ± 15.53
Mini-Zarit (score/7) (*n* = 129)	2.75 ± 1.57

*18 AD cases were confirmed by cerebrospinal fluid biomarkers, the others 103 cases were clinical probable AD (no lumbar puncture). † For the others participants, the etiologies were anxiety disorders (*n* = 2), depressive disorders (*n* = 2), vascular cognitive disorders (*n* = 4), fronto-temporal dementia (*n* = 1), cerebral amyloid angiopathy (*n* = 1), other mental disorders (*n* = 1) or still not determinate at the time of the evaluation (*n* = 7)

MMSE: Mini-Mental State Examination; IADL: Instrumental Activities of Daily Living; DAD6: 6-item Disability Assessment for Dementia; NPI: Neuropsychiatric Inventory

An alpha level of 0.05 was used for statistical significance; all tests were bilateral. Statistical analyses were performed with SPSS software (version 20, SPSS Statistics Inc.).

A sensitivity analysis was performed to ensure that the agreement between the two methods of evaluation of the CDR-SB assessed by ICC remained similar when removing from the analysis the patients who had evaluation of the CDR conducted twice at different time of their care pathway.

Results

The evaluation of the CDR with the two methods (face-to-face and with the patient's medical record) has been performed 145 times for 139 different patients, meaning that for six patients, this double CDR evaluation has been performed twice during different memory visits. The patients' characteristics are described in Table 1. There were a higher proportion of women (58.3%), and the majority of the patients had AD (87%).

Six different neuropsychologists carried out the evaluation of the CDR using the patient's medical records. The evaluation of the CDR with the patient's medical record was performed for 92 patients in the case A, for 107 patients in the case B and for 138 patients in the case C.

Agreement between face-to-face and patient's medical records methods of assessment for the CDR-SB

The ICC for the CDR-SB score assessed by face-to-face and with the patient's medical record was 0.95 (95% CI of ICC score: 0.93–0.97, excellent agreement) in the case A where all the available information of the patient's medical record was used, 0.86 (95% CI of ICC score: 0.80–0.91, good agreement) in the case B and 0.92 (95% CI of ICC score: 0.89–0.94, excellent agreement) in the case C (Table 2).

The sensitivity analysis consisting to exclude the 6 patients who were evaluated twice at different time of their care pathway showed same ICC i.e. ICC between the CDR-SB score assessed by face-to-face and the patient's medical record: 0.95 (95% CI of ICC score: 0.93–0.97) in the case A, 0.86 (95% CI of ICC score: 0.80–0.91) in the case B and 0.92 (95% CI of ICC score: 0.89–0.94) in the case C.

Using the Bland and Altman method, the mean difference between the CDR-SB score assessed by face-to-face

Table 2 Analysis of agreement between the methods of evaluation of the CDR-SB

	n	CDR-SB Mean ± SD	ICC* (95%)	Mean CDR-SB difference ± SD	Lower/Upper limits agreement	Percentage of observations inside limits agreement
Face-to-face evaluation	145	5.99 ± 2.91				
Medical record evaluation Case A†	92	5.13 ± 2.48	0.95 (0.93–0.97)	0.098 ± 1.036	(-1.974 / 2.17)	92.4%
Medical record evaluation Case B†	107	5.15 ± 2.60	0.86 (0.80–0.91)	– 0.379 ± 1.809	(-3.997 / 3.239)	94.4%
Medical record evaluation Case C†	138	5.62 ± 3.01	0.92 (0.89–0.94)	– 0.29 ± 1.583	(-3.456 / 2.876)	92.8%

* Intraclass correlation coefficient (ICC) between face-to-face Sum of boxes of the clinical dementia rating scale (CDR-SB) and CDR-SB evaluated with the patient's medical record

†Case A: all the information available in the patient's medical record is used by the evaluator to assess the CDR-SB

Case B: all the information in the patient's medical record, except the comprehensive neuropsychological assessment is used by the evaluator to assess the CDR-SB

Case C: all the information in the patient's medical record, except the comprehensive neuropsychological assessment and the medical doctor's clinical examination, is used by the evaluator to assess the CDR-SB

and with the patient's medical record in case A was 0.098±1.036, and 92.4% of the patients lay within the 95% limits of agreement (Table 2; Fig. 2). In case B, the mean difference was respectively −0.379±1.809 with 94.4% of the patients laying within the 95% limits of agreement (Fig. 2). In case C, the mean difference was respectively −0.29±1.583 with 92.8% of the patients laying within the 95% limits of agreement (Fig. 2).

Agreement between the methods of assessment for the 5-point scale CDR and the CDR sub-scores

The ICC for the 5-point scale CDR assessed by face-to-face and with the patient's medical record was 0.92 (95% CI of ICC score: 0.88–0.95, excellent agreement) in the case A where all the available information of the patient's medical record was used, 0.81 (95% CI of ICC score: 0.72–0.87, good agreement) in the case B and 0.83 (95% CI of ICC score: 0.77–0.88, good agreement) in the case C (Table 3).

The linear weighted Kappa coefficients was 0.79 (95% CI: 0.68–0.91) for the 5-point scale CDR comparison obtained by face-to-face interview and with the patient's medical record (case A) indicating a strong agreement, 0.58 (95% CI: 0.45–0.71) in the case B indicating a moderate agreement, and 0.13 (95% CI: 0.08–0.19) in the case C indicating a very weak agreement.

The analysis by domain of the CDR scale (CDR sub-scores) showed ICC ranging from 0.65 (moderate agreement) to 0.91 (excellent agreement) depending of the domains and the methods of evaluation. Regarding the linear weighted kappa coefficients, they ranged from 0.26 (weak agreement) to 0.79 (strong agreement). The level of agreement between the CDR sub-scores obtained with the patient's medical record and during face-to-face interview was better in the case A compared to cases B and C, excepted for the domain « Personal care », for which the level of agreement was slightly higher in the case C.

Description of information used in the patient's medical records to evaluate the different domains of the CDR (by domain)

The supplement Tables 2, 3 and 4 presents the proportion of use of the information in the patient's medical records to estimate the sub-scores of the different domains of the CDR.

In the case A, the evaluation of the CDR sub-scores (by domain) using all the available information of patient's medical record was mainly based on the components of the MMSE, the neuropsychology interview and assessments, as well as life context collected during nurse interview for the « Memory » domain (Supplement Table 2). It was mainly based on the orientation component of the MMSE, and the transportation and moving item of the functional scales for the « Orientation » domain ; on the MMSE components, functional items, the Frontal Assessment Battery (FAB) assessment [22], the neuropsychology interview, the occupational therapist report and the nurse interview for the « Judgement and problem solving skills » domain ; on the functional scales, and particularly the AGGIR scales for the « Personal care » domain ; on the functional scales, and more particularly the IADL, for the « Domestic activities and hobbies » domain; and on the functional scales, the neuropsychological interview, the nurse interview and the occupational therapist report for the « Activities away from home » domain.

In the case B, when the information of the neuropsychological interview and the neuropsychological assessment could not be used, the neuropsychologists used more frequently the information from the report for cognitive complaints and the NPI scale (Supplement Table 3).

In the case C, when the information of the neuropsychological interview and the neuropsychological assessment as well as the information of the medical consultation could not be used, the neuropsychologists used more frequently the information from the AGGIR scale (Supplement Table 4).

Feasibility of evaluating CDR using the patient's medical record

The comparison of the duration of evaluation of the CDR scale showed shorter duration with the patient's medical record compared to face-to-face interview (Table 4). In case A, the CDR evaluation on the patient's medical record lasted 14.65±2.89 min compared to 39.89±7.77 min for the face-to-face evaluation.

Among the 139 patients, 145 evaluations of the CDR scale were performed during face-to-face interview : 92 (63.4%) evaluations could be performed using the patient's medical records in the case A (using all the information including the neuropsychological evaluation and the clinical examination), 107 (73.8%) could be performed in the case B (all information excepted the neuropsychological evaluation), 138 (95.1%) could be performed in the case C (all information excepted the neuropsychological evaluation and the clinical examination. The lower proportion of evaluation in the case A was explained by the conditions to have a patient's medical records containing all the needed information.

Discussion

The CLIMER study allowed to study the agreement between methods of assessment of the CDR scores i.e. face-to-face interview, which represents the standard method and the analysis of the data available in the patient's medical records according to 3 cases (A : using all the information of the patient's medical record, B :

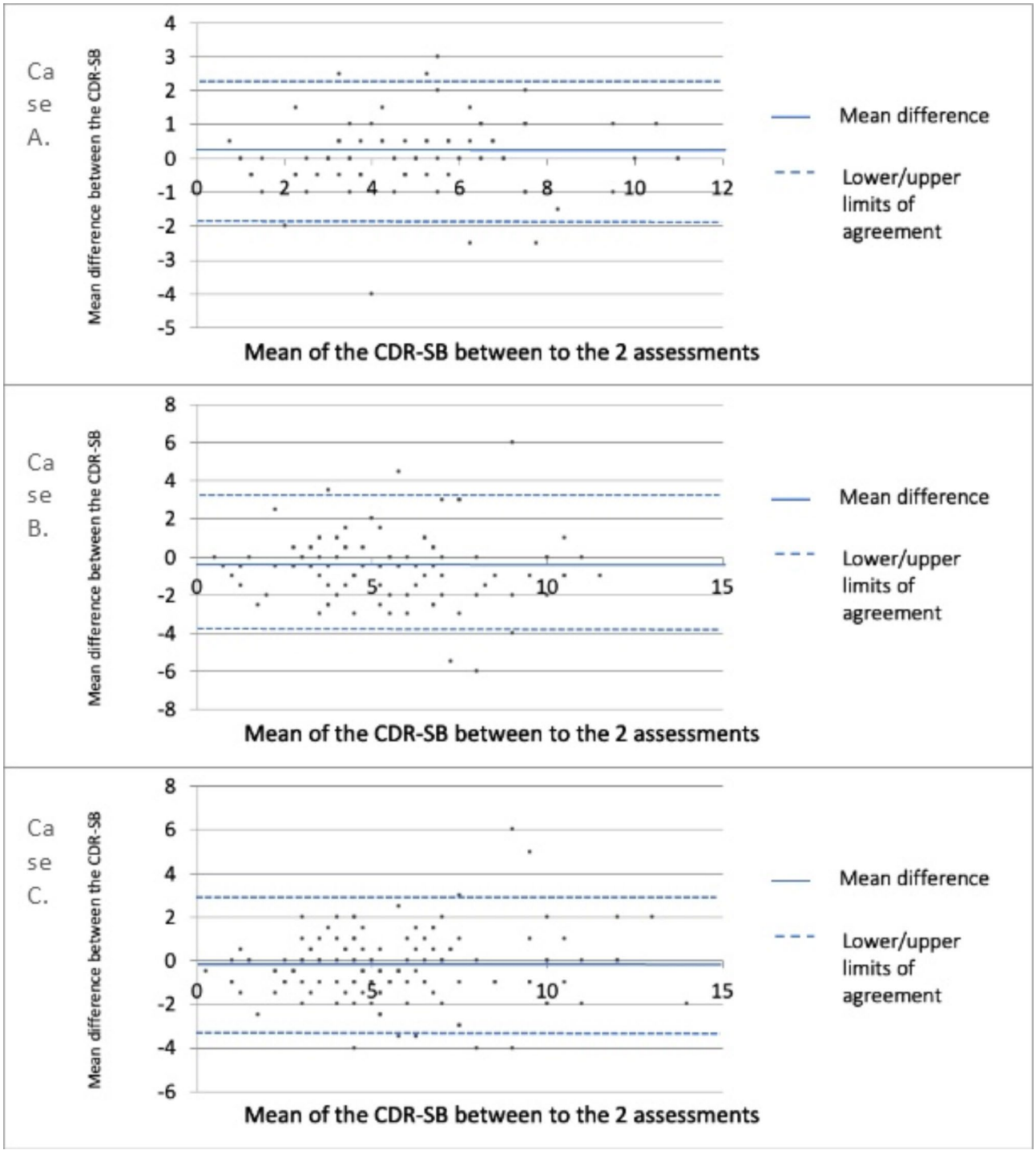

Fig. 2 Bland-Altman plots for the agreement between the CDR-SB assessed face-to-face and the CDR-SB assessed using the patient's medical record for each case. CDR-SB: Sum of boxes of the clinical dementia rating scale

using all the information excepted the neuropsychological assessments, C : using all the information excepted the neuropsychological assessments and the medical consultation including cognitive screening tests). The present study allows to highlight an excellent level of agreement of the evaluation of the CDR-SB using the all the information of the patient's medical record (case A) compared to the standard face-to-face evaluation. This result was confirmed by the Bland and Altman analysis. Another result supporting the interest of the evaluation of the CDR-SB using all the available information of the patient's medical record is that the duration of

Table 3 Description of agreement for the CDR sub-scores and the 5-point scale CDR

	n = 145	ICC* (95%)	K†	95% CI
CDR score in case A vs. face-to-face (FF)	92	0.92 (0.88–0.95)	0.79	0.68–0.91
CDR score in case B vs. FF	107	0.81 (0.72–0.87)	0.58	0.45–0.71
CDR score in case C vs. FF	138	0.83 (0.77–0.88)	0.13	0.08–0.19
Domains of the CDR score				
Memory in case A vs. FF	92	0.85 (0.76–0.90)	0.69	0.57–0.81
Memory in case B vs. FF	108	0.68 (0.53–0.78)	0.44	0.30–0.57
Memory in case C vs. FF	138	0.68 (0.55–0.77)	0.43	0.31–0.55
Orientation in case A vs. FF	92	0.87 (0.81–0.92)	0.71	0.60–0.82
Orientation in case B vs. FF	108	0.68 (0.53–0.78)	0.44	0.32–0.56
Orientation in case C vs. FF	138	0.80 (0.71–0.87)	0.57	0.47–0.67
Judgement solving problem in case A vs. FF	92	0.69 (0.53–0.80)	0.49	0.33–0.64
Judgement solving problem in case B vs. FF	107	0.66 (0.49–0.77)	0.26	0.12–0.40
Judgement solving problem in case C vs. FF	138	0.65 (0.51–0.75)	0.30	0.17–0.42
Community affairs in case A vs. FF	92	0.77 (0.65–0.85)	0.59	0.47–0.71
Community affairs in case B vs. FF	107	0.68 (0.53–0.78)	0.45	0.32–0.58
Community affairs in case C vs. FF	138	0.76 (0.66–0.83)	0.47	0.37–0.58
Home & hobbies in case A vs. FF	92	0.91 (0.87–0.94)	0.67	0.55–0.79
Home & hobbies in case B vs. FF	107	0.77 (0.66–0.84)	0.50	0.37–0.63
Home & hobbies in case C vs. FF	138	0.85 (0.79–0.89)	0.62	0.52–0.71
Personal Care in case A vs. FF	92	0.88 (0.81–0.92)	0.73	0.59–0.87
Personal Care in case B vs. FF	107	0.76 (0.64–0.83)	0.54	0.40–0.69
Personal Care in case C vs. FF	138	0.91 (0.88–0.94)	0.79	0.70–0.87

* Intraclass correlation coefficient (ICC)

† Linear weighted kappa (K)

CDR: Clinical Dementia Rating; IC: Confident Intervals

Table 4 Comparison of duration of evaluation of the CDR scale

	Duration of evaluation of the CDR scale Mean ± SD (in minutes)	*P* value*
Face-to-face interview	39.89 ± 7.77	
Patient's medical record		
Case A	14.65 ± 2.89	< 0.0001
Case B	6.09 ± 3.22	< 0.0001
Case C	5.90 ± 2.51	< 0.0001

* paired t-test to compare the mean time between each case and the face-to-face interview

CDR: Clinical Dementia Rating

the evaluation was shorter compared to the face-to-face interview (less than 37% of the time for face-to-face interview). The level of agreement was considered as good for the cases B and C, in which less information was explored to evaluate the CDR-SB in the patient's medical record, with even shorter duration of evaluation.

When studying the 5-point scale CDR, the level of agreement was still higher in the case A than in the cases B or C and considered as excellent for the evaluation with the patient's medical record in the case A based on the ICC.

The analysis of the agreement with the linear weighted Kappa coefficients showed generally less favorable results in particular in the case C. This discrepancy between these both measures of agreement the ICC and the Kappa has previously been observed and may be explained by the fact that the ICC is more adapted for quantitative variables, whereas the Kappa is more adapted for categorical variables [23].

The study of level of agreement between the different domains of the CDR-scale showed generally higher level of agreement when all the information of the patient's medical record is used (case A) than when only part of the information is used (cases B and C). Given these findings, the evaluation of the CDR-scale by neuropsychologists using all the information available in the patient's medical (case A) record appears very reliable when the face-to-face interview cannot be carried out, in patients attending a Memory consultation.

The present study confirmed and extend previous finding showing significant correlations between CDR-SB and neuropsychological tests, the MMSE, and the ADAS-Cog [8, 9, 24]. We showed that the MMSE components and the neuropsychological tests present in the patient's

medical records are information that has often been used to evaluate the CDR scale with this method. Nevertheless, previous studies limit the comparison with specific scales, and do not take into account other information that may be available in the patient's medical records. The CLIMER study is then complementary of these studies by showing that the use of patient's medical records containing various important information allows a reliable evaluation of the CDR scores, with a shorter duration of the CDR evaluation and a facilitated organization since the patients and their caregivers do not need to be interviewed.

There is limited research on the use of patient's medical records to evaluate and extrapolate an existing scales, and so for the CDR scale, without specifically administering the scale by face-to-face interview. However, in Wilhelmson et al., the authors compared the information regarding illness, symptoms and impairment in older patients, from their patient's medical records and from interviews [25]. The authors have considered that both sources did not provide the same level of information, the medical records provide better information in case of specific diseases and for the diagnoses, and the interview provide better measure of illness, functional impairment and global status health. In the case of the present study, it should be considered that the information used from the patient's medical records were collected from previous face-to-face interviews between the patient, its caregiver and different care staffs. Some information collected in previous scales could be easily extrapolated to evaluate the different domains of the CDR scale.

To our knowledge, this study was the first to assess the agreement of different methods of evaluation of the CDR scale: face-to-face interview and patient's medical records and its results highlight that the patient's medical records analysis by trained neuropsychologists provide a reliable alternative method of assessment of the CDR scale compared to face-to-face interview, and in particular the CDR-SB, in the context of real-life. This expands possibilities in the field of real-world data collection in AD, as CDR-SB constitutes a relevant scale in AD, and could be assessed based on patient's medical records.

Limitations

The evaluation of the CDR scale based on the patient's medical records contained an implicit judgement of the neuropsychologist. Questioning about the replicability of the evaluation of the CDR scale implicating implicit judgement may arise. Nevertheless, one cannot rule out the fact that implicit judgement is also actually occurring during face-to-face interview, leading to variations between the different evaluators, while the CDR scale has been shown to have a good inter-rater agreement [26]. Several neuropsychologists formed for the evaluation of the CDR scale have intervened in the study so that the different CDR assessments were carried out blind between all the different methods to avoid bias in the evaluations. Three different situations were also considered depending on the presence of the neuropsychological evaluation and the clinical evaluation by the physician, allowing to take into account different cases in routine care. In addition, the information collected in the patient's medical records to assess the CDR sub-scores was also collected and described in order to identify which information were mainly used by the neuropsychologists.

The evaluation by the patient's medical records has some limits that should be taken into account when interpreting the results. While the evaluation of the CDR scale by the patient's medical record in the case A when all the information was available has an excellent level of agreement with the face-to-face evaluation, its evaluation was less often feasible since it required to have all the required information. Indeed, the number of patients with more complete medical records was lower in the case A than in the cases B and C in which less information was needed to conduct the evaluation. Consequently, the sample size to assess the agreement between the different methods of evaluation the CDR scale was smaller in the case A. Nevertheless, this sample size remained similar to the median sample size of previous studies conducting agreement analyses i.e. median sample size of 50 for categorical outcome and median sample size of 119 for continuous outcome [27].

The reliability of the assessment using the patient's medical records was conducted for measures at equivalent time with a tolerate delay of 2 months. However, we cannot exclude that the patients' health status has changed within these 2 months, in particular for older patients with neurocognitive disorders, which may explain in part some difference between the evaluations.

The evaluations were conducted twice for six patients at different time of their care pathway. These evaluations were not expected and are explained because independent neuropsychologists performed the evaluation. We chose to keep these evaluations in the analyses and to consider them as independent, as the results of a sensitivity analysis showed same level of agreement between the evaluations when they were excluded.

In further research, an algorithm to extrapolate the CDR-SB based on explicit information available in the medical records could be calculated and in case its good performance is confirmed, it could be implemented to calculate automatically the CDR-SB. In the present study, not all the information used from the medical records could be coded, as there was an implicit evaluation of the neuropsychologists. In particular, the information collected from medical reports is not systematically electronically collected. Finally, the AD diagnosis was

confirmed by cerebrospinal fluid biomarkers for a small number of patients, as lumbar puncture is not part of the routine care.

Generalizability

The present study was conducted on patients attending a memory center, and for whom several scales are collected in routine care to evaluate the cognitive performance and functional abilities. This context offers the possibility to extrapolate information for the patient's medical record in order to evaluate the CDR scores. This method of evaluation could be generalized in similar condition from medical record containing this information. Patients attending a memory center in France will have generally similar evaluations that should allow to extrapolate the CDR scores from their medical records.

These evaluations were conducted by neuropsychologists who are trained for the evaluation of the CDR scale, and for the evaluation of patients with neurocognitive disorders. They have habits to interpret the different information in the patient's medical record and they will implicitly look for corresponding information from the different scales. We do not know whether the present results could be generalized to other medical staff whom do not have this knowledge and experience, and further research should be conduct to answer this question.

Conclusions

This study showed that the evaluation of the CDR-SB and the 5-point scale CDR using all the information of the patient's medical record had an excellent level of agreement with the evaluation of the CDR-SB during face-to-face interview, in patients attending a memory center. To a less extend the evaluation of the CDR-SB using the medical record without the neuropsychological evaluation or the medical evaluation had a good agreement with the face-to-face interview.

When the face-to-face interview cannot be performed due to limited time, the evaluation of the CDR-SB using the patient's medical record in the memory center by neuropsychologists provides a reliable evaluation of this score. Moreover, in the perspective of real-world data collection in AD, we hypothesize that CDR-SB may be assessed in several patients in large multi-center studies, even if face-to-face CDR-SB is not available.

Abbreviations

AD	Alzheimer's Disease
CDR	Clinical Dementia Rating
CDR-SB	CDR sum of boxes
DAD6	6-item Disability Assessment for Dementia
IADL	Instrumental Activities of Daily Living
ICC	Intraclass correlation coefficient
MMSE	Mini-Mental State Examination
MCI	Mild Cognitive Impairment
NCD	Neurocognitive Disorders
NPI	Neuropsychiatric Inventory

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13195-024-01567-9.

Supplementary Material 1

Acknowledgements

We are grateful to Dr Frédéric Gervais, Laurie Spirli, Romain Bachelet, Cynthia Claus and Marie Gouy, the participants and the hospital staff for their contribution to this study.

Author contributions

All authors have contributed substantially to this work and meet criteria for authorship as stated in the Uniform Requirements for Manuscripts Submitted to Biomedical Journals, and each authors' contribution is listed below: Study concept and design: Virginie Dauphinot (VD), Sylvain Calvi (SC), Anthony Bathsavanis (AB), Claire Moutet (CM), Pierre Krolak-Salmon (PKS), Antoine Garnier-Crussard (AGC). Acquisition of data: PKS, AGC, CM, SC, AB, Sophie Dautricourt (SD), Jing Xie (JX). Analysis and interpretation of data: VD, SC, AB, AGC, CM. Drafting of the manuscript: VD, SCCritical revision of the manuscript for important intellectual content: VD, SC, AB, AGC, PKS, CM, SD, JX. All authors have read and approved this final version of the manuscript.

Funding

This study is an ancillary study of the MEMORA cohort study and it was funded by BIOGEN. The MEMORA cohort and the present study are sponsored by the University Hospital of Lyon. The MEMORA cohort is supported by a grant from the MSD Avenir Foundation. Funding sources were not involved in the data acquisition, analysis, and manuscript writing.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

This study was performed in accordance with the principles of the Declaration of Helsinki. The opinion of the scientific and ethical committee of the Hospices Civils de Lyon was requested, and the favorable opinion was obtained on 12/7/2020. The CLIMER study was classified as a non-interventional research that does not involve the human person (RNIPH), and as such, signed informed consent from patients was not required under the French general regulation for data protection (RGPD). However, individual information has been sent to patients (and their caregivers) to inform them of the use of patient health data in the context of this research. Patients could object to the use of their health data. The data processing of this research was part of the MR-004 reference methodology and as such, the declaration form has been completed. The study is reported on the clinicaltrials registry: https://clinicaltrials.gov/ct2/show/NCT04763941.

Consent for publication

Not applicable.

Competing interests

Virginie Dauphinot, Sylvain Calvi, Claire Moutet, Jing Xie, Sophie Dautricourt, and Anthony Bathsavanis, Antoine Garnier-Crussard are working at the University Hospital of Lyon. Pierre Krolak-Salmon was working at University Hospital of Lyon at the time of the study. At the time of the study, and independent of this work, AGC, JX and SD are unpaid sub-investigators or principal investigator in NCT04867616 (UCB Pharma), NCT04241068 (Biogen), NCT05310071 (Biogen), NCT03446001 (TauRx Therapeutics), NCT03444870 (Roche), NCT04374253 (Roche), NCT04777396 (Novo Nordisk), NCT04777409 (Novo Nordisk), NCT04770220 (Alzheon), NCT05423522 (Medesis Pharma). Independent of this work, PKS was an unpaid sub-investigator or principal investigator in NCT04867616 (UCB Pharma), NCT04241068 (Biogen), NCT03446001 (TauRx Therapeutics), NCT03444870 (Roche), NCT04374253 (Roche), NCT04777396 (Novo Nordisk), NCT04777409 (Novo Nordisk), NCT04770220 (Alzheon), NCT05423522 (Medesis Pharma), and provides consultancy for Biogen, Roche, Novartis, MSD, Lilly, Pfizer, Abbvie.

Author details
[1]Clinical and Research Memory Center of Lyon, Lyon Institute For Aging, Charpennes Hospital, Hospices Civils de Lyon, Lyon, France
[2]PhIND "Physiopathology and Imaging of Neurological Disorders", Neuropresage Team, Normandie Univ, UNICAEN, INSERM, U1237, Cyceron, Caen 14000, France
[3]Clinical and Research Memory Center, Hôpital des Charpennes, 27 rue Gabriel Péri, Villeurbanne 69100, France

Received: 29 November 2023 / Accepted: 25 August 2024
Published online: 05 September 2024

References

1. Hughes CP, Berg L, Danziger WLCLA, Martin RL. A new clinical scale for the staging of dementia. Br J Psychiatry. 1982;140:566–72.
2. O'Bryant SE, Lacritz LH, Hall J, Waring SC, Chan W, Khodr ZG, et al. Validation of the new interpretive guidelines for the clinical dementia rating scale sum of boxes score in the national Alzheimer's coordinating center database. Arch Neurol. 2010;67(6):746–9.
3. Nyunt MS, Chong MS, Lim WS, Lee TS, Yap P, Ng TP. Reliability and validity of the clinical dementia rating for Community-Living Elderly subjects without an informant. Dement Geriatr Cogn Dis Extra. 2013;3(1):407–16.
4. Cedarbaum JM, Jaros M, Hernandez C, Coley N, Andrieu S, Grundman M, et al. Rationale for use of the clinical dementia rating Sum of boxes as a primary outcome measure for Alzheimer's disease clinical trials. Alzheimer Dement. 2013;1:S45–55.
5. Budd Haeberlein S, Aisen PS, Barkhof F, Chalkias S, Chen TF, Cohen S, et al. Two Randomized phase 3 studies of Aducanumab in Early Alzheimer's Disease. J Prev Alzheimers Dis. 2022;9(2):197–210.
6. van Dyck CH, Swanson CJ, Aisen P, Bateman RJ, Chen CC, Gee M, et al. Lecanemab in Early Alzheimer's Disease. N Engl J Med. 2023;388(1):9–21.
7. Sims JR, Zimmer JA, Evans CD, Lu M, Ardayfio P, Sparks J, et al. Donanemab in early symptomatic Alzheimer Disease: the TRAILBLAZER-ALZ 2 Randomized Clinical Trial. JAMA. 2023;330(6):512–27.
8. Fillenbaum GG, Peterson B, Morris J. Estimating the validity of the clinical dementia rating scale: the CERAD experience. Consortium to establish a Registry for Alzheimer's Disease. Aging. 1996;8(6):379–85.
9. Perneczky R, Wagenpfeil S, Komossa K, Grtimmer T, Diehl J, Kurz A. Mapping scores onto stages: Mini-mental State Examination and Clinical Dementia Rating. Am J Geriatr Psychiatry. 2006;14(2):139–44.
10. Dauphinot V, Moutet C, Rouch I, Verdurand M, Mouchoux C, Delphin-Combe F, et al. A multicenter cohort study to investigate the factors associated with functional autonomy change in patients with cognitive complaint or neurocognitive disorders: the MEMORA study protocol. BMC Geriatr. 2019;19(1):191.
11. Jessen F, Amariglio RE, Buckley R, Van der Flier WM, Han Y, Molinuevo JL, et al. The characterisation of subjective cognitive decline. Lancet Neurol. 2020;19:271–78.
12. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnosis guidelines for Alzheimer's disease. Alzheimers Dement. 2011;7(3):270–9.
13. McKhann G, Knopman D, Chertkow H, Hyman B, Jack Cj, Kawas C, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association worgroups on diagnosis guidelines for Alzheimer's disease. Alzheimer Dement. 2011;7(3):263–9.
14. O'Bryant SE, Waring SC, Cullum CM, Hall J, Lacritz LH, Massman PJ, et al. Staging dementia using clinical Dementia Rating Scale Sum of boxes scores: a Texas Alzheimer's research consortium study. Arch Neurol. 2008;65(8):1091–5.
15. Folstein M, Folstein S. Mini-mental state: a practical method for grading the cognitive stade of patients for the clinician. J Psychiatr Res. 1975;12(3):189–98.
16. Lawton M, Brody E. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969;9(3):179–86.
17. de Rotrou J, Wu YH, Hugonot-Diener L, Thomas-Antérion C, Vidal JS, Plichart M, et al. DAD-6: a 6-Item version of the Disability Assessment for Dementia scale which may differenciate Alzheimer's disease and mild cognitive impairment from controls. Dement Geriatr Cogn Disord. 2012;33(2–3):210–8.
18. Bedart M, Molloy DW, Squire L, Dubois S, Lever JA, O'Donnell M. The Zarit Burden interview: a new short version and screening version. 41. 2001;5:652–7.
19. Cummings JL. The Neuropsychiatric Inventory: assessing psychopathology in dementia patients. Neurology. 1997 May 1997;48:S10-6.
20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986;1(8476):307–10.
21. Fleiss JL, Cohen J. The equivalence of Weighted Kappa and the Intraclass correlation coefficient as measures of reliability. Educ Psychol Meas. 1973;33:613–9.
22. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: a Frontal Assessment Battery at bedside. Neurology. 2000;55(11):1621–6.
23. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43(6):543–9.
24. Balsis S, Benge JF, Lowe DA, Geraci L, Doody R. How do scores on the ADAS-Cog, MMSE, and CDR-SOB Correspond? Clin Neuropsychol. 2015;29(7):1002–9.
25. Wilhelmson K, Rubenowitz Lundin E, Andersson C, Sundh V, Waern M. Interviews or medical records, which type of data yields the best information on elderly people's health status? Aging Clin Exp Res. 2006;18(1):25–33.
26. Tractenberg RE, Schafer K, Morris JC. Interobserver disagreements on clinical dementia rating assessment: interpretation and implications for training. Alzheimer Dis Assoc Disord. 2001;15(3):155–61.
27. Han O, Wei Tan H, Julious S, Sutton L, Jacques R, Lee EK, et al. A descriptive study of samples sizes used in agreement studies published in the PubMed repository. BMC Med Res Methodol. 2022;22:242.

Publisher's note