ELSEVIER

Alzheimer's & Dementia: Translational Research & Clinical Interventions ■ (2018) 1-10

Alzheimer's & Dementia

Featured Article

TOMMORROW neuropsychological battery: German language validation and normative study

Heather R. Romero[a,b], Andreas U. Monsch[c], Kathleen M. Hayden[a,b,d], Brenda L. Plassman[a,b], Alexandra S. Atkins[e], Richard S. E. Keefe[b,e], Shyama Brewster[f], Carl Chiang[f], Janet O'Neil[g], Grant Runyan[g], Mark J. Atkinson[h,i], Stephen Crawford[h], Kumar Budur[g], Daniel K. Burns[f], Kathleen A. Welsh-Bohmer[a,b,*], for the TOMMORROW Study Investigators

[a]*Joseph & Kathleen Bryan Alzheimer's Disease Research Center (Bryan ADRC), Duke University, Durham, NC, USA*
[b]*Department of Psychiatry, Duke University, Durham, NC, USA*
[c]*University Center for Medicine of Aging, Felix Platter Hospital, Basel, Switzerland*
[d]*Department of Social Sciences and Health Policy, Wake Forest School of Medicine, Winston-Salem, NC, USA*
[e]*NeuroCog Trials, Durham, NC, USA*
[f]*Zinfandel Pharmaceuticals, Inc., Chapel Hill, NC, USA*
[g]*Takeda Development Center, Americas, Inc., Deerfield, IL, USA*
[h]*Covance Inc., Princeton, NJ, USA*
[i]*Department of Family Medicine and Public Health, University of California, San Diego, CA, USA*

Abstract

Introduction: Assessment of preclinical Alzheimer's disease (AD) requires reliable and validated methods to detect subtle cognitive changes. The battery of standardized cognitive assessments that is used for diagnostic criteria for mild cognitive impairment due to AD in the TOMMORROW study have only been fully validated in English-speaking countries. We conducted a validation and normative study of the German language version of the TOMMORROW neuropsychological test battery, which tests episodic memory, language, visuospatial ability, executive function, and attention.

Methods: German-speaking cognitively healthy controls (NCs) and subjects with AD were recruited from a memory clinic at a Swiss medical center. Construct validity, test–retest, and alternate form reliability were assessed in NCs. Criterion and discriminant validities of the cognitive measures

Conflicts of interest: H.R.R., K.M.H., B.L.P., and K.A.W.-B. received funding from Takeda as part of a contract with Duke University for the work she and her team conducted as the neuropsychology leads to the TOMMORROW program. A.U.M. has received funding from Takeda/Zinfandel to conduct the study. He currently or in the last 3 years has received honoraria and served as a consultant or advisory board member for AbbVie, AC Immune, Lilly, Roche, and Vifor Pharma. A.S.A. is an employee of NeuroCog Trials, which provided services for this study. Currently or in the past 3 years she has received funding from the National Institute of Mental Health and the National Institute on Aging. R.S.E.K. currently or in the past 3 years has received investigator-initiated research funding support from the Department of Veteran's Affairs, National Institute of Mental Health, and the Singapore National Medical Research Council. He, currently or in the past 3 years, has received honoraria, served as a consultant, speaker, or advisory board member for AbbVie, Acadia Pharmaceuticals, Aeglea BioTherapeutics, Akebia Therapeutics, Akili, Alkermes, ArmaGen, Astellas Pharma, Avanir Pharmaceuticals, AviNeuro/ChemRar, Axovant Sciences, Biogen, Boehringer Ingelheim, Cerecor, CoMentis, Critical Path Institute, FORUM Pharmaceuticals, Global Medical Education (GME), GW Pharmaceuticals, Intra-Cellular Therapeutics, Janssen, Lundbeck, Lysogene, Medscape, Merck, Minerva Neurosciences, Mitsubishi, Monteris Medical, Moscow Research Institute of Psychiatry, Neuralstem, Neuronix, Novartis, New York State Office of Mental Health, Otsuka, Pfizer, Regenix Bio, Reviva Labs, Roche, Sangamo Therapeutics, Sanofi, Sunovion Pharmaceuticals, Takeda, Targacept, University of Moscow, University of Texas Southwestern Medical Center, and WebMD. R.S.E.K. receives royalties from the BACS testing battery, the MATRICS Battery (BACS Symbol Coding, and the Virtual Reality Functional Capacity Assessment Tool (VRFCAT). He is also a shareholder in NeuroCog Trials, which provided services for this study, and SenGenix. S.B. is an employee of Zinfandel Pharmaceuticals. C.C. is an employee of Zinfandel Pharmaceuticals. J.O'N. is an employee of Takeda. G.R. was employed by Takeda at the time of this study. M.J.A. was employed by Covance at the time of the study, with academic appointment at the University of California, San Diego, USA. S.C. is an employee of Covance. K.B. was employed by Takeda at the time of this study. D.K.B. is an employee of Zinfandel Pharmaceuticals.

*Corresponding author. Tel.: +1 919 668 1553; Fax: +1 919 669 0828.
E-mail address: kwe@duke.edu

https://doi.org/10.1016/j.trci.2018.06.009

were tested using logistic regression and discriminant analysis. Cross-cultural equivalency of performance of the German language tests was compared with English language tests.

Results: A total of 198 NCs and 25 subjects with AD (aged 65–88 years) were analyzed. All German language tests discriminated NCs from persons with AD. Episodic memory tests had the highest potential to discriminate with almost twice the predictive power of any other domain. Test–retest reliability of the test battery was adequate, and alternate form reliability for episodic memory tests was supported. For most tests, age was a significant predictor of group effect sizes; therefore, normative data were stratified by age. Validity and reliability results were similar to those in the published US cognitive testing literature.

Discussion: This study establishes the reliability and validity of the German language TOMMORROW test battery, which performed similarly to the English language tests. Some variations in test performance underscore the importance of regional normative values. The German language battery and normative data will improve the precision of measuring cognition and diagnosing incident mild cognitive impairment due to AD in clinical settings in German-speaking countries.

Keywords: Alzheimer's disease; Randomized clinical trial; Neuropsychology; Cross-cultural; Validation

1. Background

Measurement of cognitive change in Alzheimer's disease (AD) trials has typically relied on well-validated metrics, such as the Alzheimer's Disease Assessment Scale-Cognitive Subscale test, a composite measure that detects treatment response across different trial settings and cultures [1]. However, as more research started focusing on the preclinical stages of AD, more sensitive metrics are required for the detection of very subtle cognitive changes that occur in the very early stages of incident mild cognitive impairment due to AD (MCI-AD). An increasing number of secondary prevention studies are now underway that are designed to test therapeutic agents that may either delay or prevent the onset of early clinical symptoms of AD [2]. These studies have selected composite end points capable of detecting cognitive decline in older adults at risk of developing AD, based on neuropsychological data from longitudinal studies of clinically healthy populations [3,4].

The TOMMORROW study is among these secondary prevention studies and examines the efficacy of medicines to delay the onset of diagnosable MCI-AD [5,6]. Unlike the other secondary prevention trials, the primary outcome is a clinically definable event, MCI, rather than a cognitive composite. The trial uses a time-to-event design, with the primary end point event defined as an adjudicated clinical diagnosis of MCI-AD based on the National Institute on Aging–Alzheimer's Association criteria [7]. This end point choice has advantages in being a clinically meaningful outcome but also presents some complexities when applied across different languages and cultures. First, the MCI-AD criteria had not yet been used as an end point in clinical trials, and the criteria needed to be operationalized to allow diagnostic standardization across multiple sites and clinicians. Second, because the diagnostic criteria are to be used in a global context, cross-cultural validation and standardization of the metrics is essential to ensure that they capture the cognitive and functional features of early-stage AD across all the cultural settings.

In the TOMMORROW study, the core clinical criteria for MCI-AD were operationalized. These criteria were made available for public review before study launch [6] and are defined as a sustained decline in both function and neuropsychological performances. Functional decline is operationalized as a change in the Clinical Dementia Rating Scale (CDR) score from 0 to 0.5, and neuropsychological performance is measured by two domains of cognition (one of which must have been episodic memory) falling below baseline performance (to at least 1.3 standard deviation [SD] below the age-adjusted normative mean) or an isolated decline from baseline in episodic memory (i.e., to at least 1.5 SD below the age-adjusted normative mean). To meet the operationalized definition of MCI-AD, these criteria are expected to be met at two consecutive study visits, approximately 6 months apart.

Criteria for MCI-AD can be applied readily in English-speaking countries using standardized cognitive assessments that are normed for the population and well validated for clinical use in cognitively healthy older adults. However, the use of these tools for diagnostic purposes in other cultures requires establishing their test reliability, validity, and normative ranges appropriate to the applicable culture. The purpose of the present study was to psychometrically validate the German language adaptation of the cognitive tests for the TOMMORROW study for application among German-speaking adult populations aged 65–88 years (the target age range for the TOMMORROW study) and to develop normative data for the battery of tests.

2. Methods

We conducted a validation and normative study of the German language version of the TOMMORROW study's neuropsychological test battery. The primary objectives of