ELSEVIER

Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 1 (2015) 103–111

Alzheimer's & Dementia

Practice effects due to serial cognitive assessment: Implications for preclinical Alzheimer's disease randomized controlled trials

Terry E. Goldberg[a,*], Philip D. Harvey[b,c], Keith A. Wesnes[d], Peter J. Snyder[e], Lon S. Schneider[f]

[a]*Litwin Zucker Center for the Study of Alzheimer's Disease, Feinstein Institute, Hofstra North Shore LIJ School of Medicine, Manhasset, NY, USA*
[b]*Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL, USA*
[c]*Research Service, Miami VA Healthcare System, Miami, FL, USA*
[d]*Wesnes Cognition, Streatley on Thames, UK*
[e]*Department of Neurology, Alpert Medical School of Brown University & Rhode Island Hospital, Providence, RI, USA*
[f]*Departments of Psychiatry, Neurology, and Gerontology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA*

Abstract

Introduction: Practice effects are characteristic of nearly all standard cognitive tasks when repeated during serial assessments and are frequently important confounders in clinical trials.

Methods: We summarize evidence that gains in neuropsychological test performance scores associated with practice effects occur as artifactual changes associated with serial testing within clinical trials. We identify and emphasize such gains in older, non–cognitively impaired individuals and estimate an effect size of 0.25 for composite cognitive measures in older populations assessed three times in a 6- to 12-month period.

Results: We identified three complementary approaches that can be used to attenuate practice effects: (1) massed practice in a prebaseline period to reduce task familiarity effects; (2) tests designed to reduce practice-related gains so that item-specific driven improvements are minimized by using tasks that minimize strategy and/or maximize interitem interference; and (3) well-matched alternate forms.

Discussion: We have drawn attention to and increased awareness of practice effect–related gains that could result in type 1 or type 2 errors in trials. Successfully managing practice effects will eliminate a large source of error and reduce the likelihood of misinterpretation of clinical trials outcomes.

Keywords: Alzheimer's disease; Practice effects; Cognition; Clinical trials; Serial assessment; Preclinical Alzheimer's disease; Neuropsychology

1. Introduction

Practice effects are characteristic of serial neurocognitive assessments, including those used in clinical trials. They refer to changes in test performance attributed to increasing familiarity with and exposure to test instruments, paradigms, and items. Nevertheless, these effects are often underappreciated. Our own work in this area [1–3] has identified them as important in the interpretation of both outcomes in clinical trials and in longitudinal studies of patients with schizophrenia. Here, we discuss the relevance of these findings to clinical trials for Alzheimer's disease (AD) and mild cognitive impairment (MCI), a stage often thought to be transitional between cognitive health and AD, and, notably, preclinical AD [4]. Preclinical AD at stages 1 and 2 refers to those individuals who have cerebrospinal fluid or positron emission tomography evidence of amyloid-β abnormalities and/or "downstream" neurodegeneration but do not demonstrate cognitive changes; at stage 3, individuals additionally suffer from subtle cognitive changes. For preclinical AD, the assessment of cognition has been suggested

*Corresponding author. Tel.: +1-516-562-0410; Fax: +1-516-562-0401.
E-mail address: tgoldber@nshs.edu

http://dx.doi.org/10.1016/j.dadm.2014.11.003

by the Food and Drug Administration (FDA) as a suitable and sole primary end point for the accelerated approval of a pharmaceutical treatment (FDA Draft Guidelines for Early Stage AD) [5]. For recent clinical trials in AD and MCI, studies typically used designs comparing cognition between the drug and placebo groups, assessed on several occasions but within a relatively short period of 18 months to 2 years, and with the end point or final assessment used as the outcome. That end point, however, may be strongly influenced by previous testing as we show in Sections 3 and 7 below. Thus, the serial testing used in these clinical trials may result in unappreciated but artifactual gains across a range of neuropsychological measures, including speed of processing, episodic memory, executive function, and working memory.

Practice effects may result from several different factors and in our view can be divided into two components. The first can be termed task familiarity and occurs early in serial assessment with given cognitive tasks. It involves the subject gaining full comprehension of the directions for the task necessary for context memory (e.g., that letters and numbers alternate in Trail-Making Test B), some knowledge of the sequence of a task (e.g., that multiple trials of a word list will be administered), and stimulus response mapping (e.g., use of a response pad in an N back test). Some task familiarity effects may be due to procedural learning, an aspect of cognition that remains relatively uncompromised in AD [6]. Even if the active treatment outperforms the placebo when both arms show practice effects, this effect may be due to an enhancement of procedural memory, which will not generate substantial benefit to the everyday cognitive function of patients with AD [7]. The second component can be termed practice-related effects. These include gains made over multiple exposures to the test because of familiarity with specific items (e.g., words on a list, a story to be recalled). Developing strategies over time that alter performance (e.g., clustering words semantically on a verbal list-learning test) might occur either as a task familiarity phenomenon or as a practice-related phenomenon. The distinction between these two components is important beyond nomenclature because it directly suggests different trial design and test construction strategies for their reduction (see Sections 3 and 7 below). If not managed, these practice effects could result in improvements that are unrelated to valid drug-placebo differences in a clinical trial.

In the context of learning and memory, practice effects would not be valid indices of specific cognitive enhancement if they do not generalize or transfer readily to other tasks or real-world activities that draw on ostensibly similar cognitive operations [8]. This is often referred to as the "transfer of training" problem. Thus, practice effects that do not relate to concurrent improvements in broad domains of cognition may be viewed as item or paradigm specific. They may also engage different cognitive operations and neural systems (e.g., procedural learning) than those thought to be treated in the intervention [9]. Also, some studies have shown that improvements in performance with repeated exposure can be used as prognostic indicators, including those related to MCI to AD conversion [10] and survival [11,12]. However, detailed discussion of these is outside the scope of the present article, which focuses on the adverse impacts of practice effects on clinical trial outcomes. Rather, in the context of a clinical trial we will cover in detail the interpretative and statistical problems associated with practice effects (see especially Sections 5.4 and 7).

We begin with a selective review of the literature on practice effects in AD, MCI, and older healthy controls as they relate to trials. We then present an example of how practice effects were confounded with treatment effects from the schizophrenia literature. Based on the literature and the schizophrenia studies, which strongly suggest that practice effects are present and large enough to obscure or be mistaken for a treatment signal, we first discuss an array of possible solutions. Next, we make recommendations for managing practice effects in preclinical AD trials based both on our review and experience in the psychometrics of test construction. It is important from the outset to recognize that our purpose is not to review the practice effect literature comprehensively. This has already been done [10,13]. Rather, our purpose is to draw out the confounding implications of practice effects in clinical trials in non–cognitively impaired older populations and suggest concrete remedies.

2. Methods

We first selectively review the literature in MCI and AD with the intention of demonstrating that even in presumptively amnestic subjects, practice effects can be identified in some cohorts. Our review in the AD and MCI groups is not meant to be exhaustive or comprehensive but rather to suggest that such effects are plausible occurrences. We then shift our focus to older, cognitively healthy individuals to demonstrate that such effects are common and measureable in serial assessment paradigms and to determine the approximate magnitude of practice effects on cognitive tests in this group. This latter group will be the focus of intense interest as the AD field moves toward secondary prevention trials in the preclinical AD spectrum.

3. Results

3.1. AD and MCI samples

Practice effects in AD have not been discussed often. Perhaps, this is the result of an expectation that many patients are substantially amnestic and unable to learn and consolidate item-level information over repeated testing. However, memory impairments are dependent on individual differences in premorbid ability and disease stage, thus creating some variability in training. Furthermore, impairments in