Validity, Significance, Strengths, Limitations, and Evidentiary Value of Real-World Clinical Data for Combination Therapy in Alzheimer’s Disease: Comparison of Efficacy and Effectiveness Studies

Alireza Atri\textsuperscript{a–c}  Susan D. Rountree\textsuperscript{d}  Oscar L. Lopez\textsuperscript{e}  Rachelle S. Doody\textsuperscript{d}

\textsuperscript{a}Department of Neurology, Massachusetts General Hospital, Boston, Mass., \textsuperscript{b}Geriatric Research, Education and Clinical Center, ENRM VA Medical Center, Bedford, Mass., \textsuperscript{c}Harvard Medical School, Boston, Mass., \textsuperscript{d}Alzheimer’s Disease and Memory Disorders Center, Department of Neurology, Baylor College of Medicine, Houston, Tex., and \textsuperscript{e}Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, Pa., USA

Key Words
Comparative effectiveness · Evidence grade · Dementia treatment · Donepezil · Galantamine · Rivastigmine · Memantine · Observational trial

Abstract
Background: Randomized controlled efficacy trials (RCTs), the scientific gold standard, are required for regulatory approval of Alzheimer’s disease (AD) interventions, yet provide limited information regarding real-world therapeutic effectiveness. Objective: To compare the nature of evidence regarding the combination of approved AD treatments from RCTs versus long-term observational controlled studies (LTOCs). Methods: Comparisons of strengths, limitations, and evidence level for monotherapy [cholinesterase inhibitor (ChEI) or memantine] and combination therapy (ChEI + memantine) in RCTs versus LTOCs. Results: RCTs examined highly selected populations over months. LTOCs collected data across multiple AD stages in large populations over many years. RCTs and LTOCs show similar patterns favoring combination over monotherapy over placebo/no treatment. Long-term combination therapy compared to monotherapy reduced cognitive and functional decline and delayed time to nursing home admission. Persistent treatment was associated with slower decline. While LTOCs used control groups, adjusted for multiple covariates, had higher external validity, and favorable ethical, practical and cost considerations, their limitations included potential selection bias due to lack of placebo comparisons and randomization. Conclusions: Naturalistic LTOCs provide complementary long-term level II evidence to complement level I evidence from short-term RCTs regarding therapeutic effectiveness in AD that may otherwise be unobtainable. A coordinated strategy/consortium to pool LTOC data from multiple centers to estimate long-term comparative effectiveness, risks/benefits, and costs of AD treatments is needed.

Introduction
Alzheimer’s disease (AD) is a chronic and progressive disease with a course of illness that spans many years, and in some individuals can stretch to more than a decade. Randomized controlled trials (RCTs) serve as the scientific gold standard for therapeutic efficacy and are required for regulatory approval of AD interventions, but are not without limitations. RCTs are often short-term studies performed in leveraged populations that cannot...
adequately inform about the long-term safety, risk-benefit calculus, and comparative costs of real-world clinical treatments [1, 2].

Long-term naturalistic observational cohort (LTOC) studies supply level II grade evidence and can provide critical information regarding effects of treating typical patients under conditions of usual clinical care; however, they are undervalued in AD for not providing randomized level I (RCT) grade evidence [3, 4]. This dilemma in clinical medicine and health policy is neither new nor unique to AD, nor was it novel in 1967 when Schwartz and Lellouch [5] elegantly analyzed these approaches and stated that ‘most therapeutic trials are inadequately formulated, and this from their earliest stages of conception … in that the trials may be aimed at the solution of one or other of two radically different kinds of problem’. They made a distinction between two different and complementary conditions and approaches, the explanatory versus pragmatic trial. The former is performed under ‘equalized’ and ‘optimized’ laboratory conditions, while the latter is performed under ‘normal’ and ‘practical’ conditions. This brief paper presents an overview of the nature of evidence with respect to comparative strengths, limitations and evidence levels for RCTs versus LTOCs for the only FDA-approved treatments of AD, cholinesterase inhibitors (ChEIs) and memantine.

**Methods**

We compared the evidentiary level for ChEI and/or memantine treatment studies including systematic database reviews/meta-analyses, RCTs, open-label extensions of RCTs (RCTOLEX) and observational cohort studies.

**Results**

RCTs across the AD severity spectrum provide level I (highest grade) evidence for 24- to 28-week on-label (FDA indication) stage-appropriate treatment efficacy and safety of ChEIs and memantine as mono- or combination therapy in AD. While the majority of these RCTs evaluated efficacy of ChEIs or memantine monotherapy [6–12], several have assessed efficacy of combination therapy (ChEI + memantine) [13, 14]. Limitations of these RCTs include their performance under idealized conditions in highly selected samples with strict inclusion/exclusion, treatment adherence and monitoring criteria, and relatively short durations (approx. 6 months, except for two 52-week RCTs with donepezil [15, 16]) relative to the course of AD dementia (approx. 5–15+ years). Despite providing further support for cognitive and functional benefits of sustained treatment with a ChEI over several years, a controversial community-based long-term RCT (AD 2000) [17] was hampered by design flaws (e.g. several on-off titrations, site-level dropout, high likelihood of selection bias, underpowered) and excessive attrition (>97% at year 3) making interpretation problematic, and the study failed to provide level I grade evidence. In response, the ongoing DOMINO-AD study [18], a long-term RCT, was launched. An important gap in the data that is to be assessed by the DOMINO-AD study is the comparison of combination therapy versus memantine alone (as well as vs. ChEI alone and vs. placebo); results from this landmark UK study are highly anticipated.

RCTOLEX studies provide level II-3 (level IIb/c) support for the benefits of sustained treatment with ChEIs for up to several years [19–22] and with memantine for 24 weeks [23]. RCTOLEX studies have additional limitations that include potential confounds due to possible effects of differential attrition, unblinding, and the absence of a control group for comparison during the open-label phase [24].

Several systematic reviews/meta-analyses, including from the Cochrane database, provide level I evidence for the short-term, on-label, and stage-appropriate efficacy of ChEIs and/or memantine in AD; these include evidence of the benefits of ChEIs [25, 26] and memantine [27] in cognition, the benefits of memantine in global severity, cognition, function, and behavior [28], and quantification of small to medium standardized effect size estimates, Cohen’s d of approximately 0.1–0.4, that favor ChEI treatment over placebo [29, 30], and memantine or ChEI + memantine combination over placebo [28] in AD.

In comparison with data from the above studies, prospectively collected data from several naturalistic LTOC studies of patients with AD treated in the clinical setting provide level II-1/2 (level IIa/b) evidence for the effectiveness of ChEIs and combination therapy (ChEI + memantine) [31, 35]. Relative to RTC populations, LTOC studies include subjects with significant comorbidities, greater concomitant and neuropsychoactive medications, and imperfect treatment adherence. Naturalistic LTOC studies also include larger populations (sample size of approx. 380–950), they have no strict inclusion/exclusion criteria (other than clinical criteria for AD), longer durations (mean follow-up of approx. 2.5–4 years), assessment of measures other than those required for regulatory approval, and gather data over multiple AD stages. Overall,
these studies provide evidence for sustained clinical effectiveness of ChEIs [31, 34, 35] and combination therapy [31], with superiority of combination therapy over time in reducing decline in cognition (fig. 1a) and function (fig. 1b) [31], delaying time to nursing home placement [32], and providing more benefit with greater persistence in treatment [31, 33]. Cohen’s d effect size estimates for treatment benefits of combination therapy versus ChEI monotherapy range from 0.1 to 0.5 for cognition and from 0.2 to 0.7 for daily function and increase over 4 years [31]. Limitations of these LTOC data include lack of randomization, use of control cohorts without the same start/stop dates as treatment cohorts, and lack of memantine monotherapy cohorts.

**Discussion**

'Most real problems contain both explanatory and pragmatic elements, for ethical reasons. Most trials hitherto have adopted the explanatory approach without question; the pragmatic approach would have often been more justifiable' [5]. Naturalistic LTOC studies provide this type of useful pragmatic data to assess long-term treatment effects in AD. Naturalistic LTOC studies in AD provide level II grade evidence, have a high external, ecological, content and convergent validity, and support data from short-term RCTs that ChEI and memantine combination is superior to ChEI alone in reducing long-term cognitive and functional decline, and delaying nursing home admission. Furthermore, LTOC studies support that benefits of ChEIs and memantine increase with persistence in therapy and over time. While short-term explanatory RCTs in AD remain a gold standard, they are a means to drug approval, not the end of assessment and discussion regarding treatment benefits and safety. Ethical considerations must take precedence over trial design, practical considerations will follow; as such, LTOC studies provide important information that may otherwise be impractical to obtain.

LTOC data complement data from RCTs and can iteratively inform their design and implementation.
Strengths of LTOC data relative to RCTs include high validity, lower cost, greater power (larger sample sizes and durations), better capacity to assess questions of practical importance, and ethical and practical favorability. Limitations of LTOC data include the fact that drug therapy is often not assigned randomly, so uncontrolled variables related to disease severity and/or duration of symptoms, perceived benefits from treatment, or other unmeasured factors may influence observed relationships between drug use and outcomes; that selection factors associated with long-term drug exposure cannot be ruled out as alternative explanations for findings, and that models can adjust for baseline differences and interactions of baseline variables with time, yet these adjustments may not fully control for potential confounding factors.

In summary, naturalistic LTOC studies are important, provide ancillary support, and should be utilized in the long-term process of therapeutic discovery and assessment including monitoring (pharmacovigilance) of long-term risk-benefit calculus to society. The integration of databases that can examine large cohorts of treated AD patients will provide invaluable information about the clinical trajectory of AD, improve prognostication, and inform clinical RCT design and therapeutic efforts in AD.

Acknowledgements

This study was supported by NIA K23AG027171 (Atri) and NIA P50AG05133 (Lopez).

References


