

Improving Screen Failure and Recruitment Rates in Alzheimer's Disease Clinical Trials



The failures of clinical trials in Alzheimer's Disease (AD) have been attributed to a variety of factors, including an inadequate understanding of mechanisms of action and/or poor target engagement; however, other factors such as inadequate study design, wrong clinical stage of AD matched to the drug's mechanism of action, limited statistical power of endpoint measures, and inclusion of participants who may otherwise not be eligible for the trial, have all contributed to the poor success rate of AD trials. In fact, failure to meet entry criteria in randomised controlled clinical studies in AD focusing on cognition improvement/sparing is a fundamental aspect in the study execution process, leading to protracted timelines and dramatically increased study costs. The importance of appropriate study design and optimisation of recruitment/screen fail rates are especially important as the field moves toward studies of putative disease-modifying agents of AD and patients that are very early in the disease spectrum – studies that have notoriously high screen failure rates (with averages upwards of 85%) and correspondingly low recruitment rates (with averages of 0.19 patients per site per month or less).

Challenges to Recruitment

Studies examining the rates of patient eligibility have established that as little as 10–27% of potential AD patients are actually trial-eligible.^{1,2} Unfortunately, only a portion of AD patients are even marginally aware of research opportunities and many are unable or simply unwilling to participate. Many older adults live alone and may not have access to a caregiver who can accompany them to study visits and aid with various procedures. Indeed, AD trials require not one but two participants: the patient and a study partner, and enrolment of this dyad is imperative to clinical trial success.

Of interest, substantive differences have been noted between enrolled AD samples and the general AD population which reflects the often idiosyncratic subject entry/eligibility criteria specific to any given study. More often than not, the diagnosis of AD in clinical practice is based on the individual clinician's distinctive diagnostic approach rather than specific research criteria. In fact, the greatest challenge for most investigators is how to properly select the right patients for a particular AD study and appropriately translate that patient's medical data and history into protocol-specific entry criteria. This becomes even more important in oligosymptomatic disease presentation in early or prodromal AD, where a patient's spontaneous reports of memory impairment are very often rare, inconsistent, and oftentimes have not been taken seriously.

AD trial recruitment is challenging due to many factors, including medical comorbidities, extensive use

of prescribed and OTC medications, and behavioural complications of AD which can all be exclusionary. Additionally, some AD patients are anxious about lumbar puncture for cerebrospinal fluid examinations or MRI/PET imaging procedures, whereas others might have difficulties with extensive and numerous psychometric tests, which often require between three to five hours to complete, and can result in frustration and emotional anguish upon confrontation of deficits.

Historical Reasons for Screen Fail Rates in AD

The development of symptomatic treatment in mild to moderate AD has traditionally been associated with average screen failure rates ranging between 15–35% in registration clinical trials. Although this range is mostly viewed as manageable by sponsors and CROs, it is not uncommon for trials to have twice the screen fail rates in early AD populations. For example, in a study of early AD patients, screen failure rates have reportedly upwards of 50%.³ One reason for this higher-than-expected screen fail rates stems from amyloid-related imaging abnormality (ARIA) exclusion criteria, stemming from the failed AN-1792 trial in which dosing in a 372-patient, multinational Phase IIa trial in patients with mild to moderate AD was suspended when four treated AD patients developed brain inflammation that later was demonstrated to reflect aseptic meningoencephalitis. In addition, this clinical trial programme established procedures that were instituted subsequently in numerous other immunotherapy programmes, and even in studies with dissimilar mechanisms of action, such as the practice of utilising a central reader to assess ARIA at baseline and at regular intervals throughout the trial. Although the original FDA guidance directed clinical trial sponsors to exclude participants with more than two existing brain microhaemorrhages from studies, the Alzheimer's Association working group proposed that research participants with up to four pre-existing microhaemorrhages (or ARIA-H) could enrol in clinical trials after reviewing all publicly available data. Additionally, any patient who develops oedema (or ARIA-E) during the trial must be taken off medication until those complications clear, and then treatment can resume. Any patients developing ARIA-H during the trial may continue to receive treatment, provided that these abnormalities do not worsen symptoms. Because microhaemorrhages cannot be easily seen in routine diagnostic sequences of brain MRI for AD, the additional MRI scans with specific sequences need to be conducted to exclude patients with ARIA-H, resulting in higher screen failure rates (i.e. average rate of 63%).⁴

Screen Failure Rates in Prodromal AD/MCI due to AD

Biological substrates of AD can be identified long before patients exhibit clinical signs and symptoms, permitting

clinical trials to enrol patients very early in the disease state, and well in advance of fulfilment of criteria for dementia. Diagnostic criteria for prodromal and/or mild cognitive impairment (MCI) due to AD have been developed which include evidence of amyloid burden and/or neurodegeneration. Although amyloid PET scanning or CSF amyloid measurement is integral for identifying subjects who are highly likely to develop AD, the expense and relative limited availability of PET scanners uniformly throughout various regions, and the regional difficulties of obtaining lumbar punctures, limits their widespread application for most multinational AD trials.

Of note, reduced CSF A β 1–42 may represent a relatively earlier biomarker than increased brain amyloid-beta detected by PET, whereas amyloid PET may be a more sensitive marker of disease progression during the development of Alzheimer's disease pathology. In fact, A β 1–42 and tau (T-tau or P-tau) are seen as relatively superior biomarkers when used in combination, as evidenced by the CSF AD signature, which combines low A β 1–42 and high T-tau or P-tau concentrations, and significantly increases the accuracy of eventual AD conversion, even at a prodromal stage. This accuracy in spite of concerns regarding the large variability in CSF methodologies between laboratories and across techniques, and the lack of agreement on CSF thresholds for inclusion purposes.

However, neither amyloid PET nor CSF biomarkers can be used as standalone tests, and should be interpreted in a larger clinical context with clear evidence of progressive memory impairment. For instance, approximately 15% of patients with clinically diagnosed AD and 34% of patients with amnesic MCI enrolled in the ADNI cohorts were quantitatively negative on amyloid PET using florbetapir⁵. The inclusion of non-amyloid burdened individuals in clinical trials of amyloid-modifying drug therapies could result in exposure to a treatment with no potential benefit and reduce the statistical power of the trial, rendering the observation of successful amyloid-modifying treatment less likely. Such screening

for amyloid positivity is now a routine part of enrolment criteria in clinical trials, but this one criteria effectively eliminates approximately 1 out of 7 patients with AD and 1 out of 3 patients with MCI, thereby preventing administration of amyloid beta-modifying treatment to patients without amyloid pathology⁵. Furthermore, the presence of reduced CSF A β 1–42 was a key entry criterion in two recent studies in patients with prodromal AD (avagacestat in prodromal AD-NCT00890890; and gantenerumab in prodromal AD-NCT01224106). In both of these studies, this particular exclusion criterion was also reportedly the main cause of extremely high screen failure rate, of approximately 80%.

In order to partially ameliorate such high screen fail rates, a hierarchic approach to patient's eligibility factors should be utilised, that takes into account all known and estimated screening variables. This hierarchy should be based on how costly and cumbersome various screening procedures are, with less costly and complex procedures occurring before others. For example, initially a subtle memory impairment with no dementia should be confirmed by specific protocol defined neuropsychological tests, whereas the lack of AD and dementia should be confirmed by global CDR score 0.5. If positive, the screening process then proceeds with safety laboratory and ECG examinations. Eventually, if these elements are favourable, a structural diagnostic and safety (ARIA-H) MRI imaging will be performed. If all examinations performed pass inclusion/exclusion criteria, then and only then should amyloid burden AD be confirmed, either by CSF examination or amyloid PET, subject to protocol. Following such a hierarchical procedure has been shown to reduce the screen failure rate in an ongoing study in prodromal AD patients to well below the expected rate, from 80% to less than 50%.

Predictable vs Unpredictable Causes of Screen Failure

We have also found it useful to conceptualise reasons for screen failure as falling into three main categories; those that are completely predictable to high quality sites such as age, medical history, progression of



cognitive decline and medical status; those that are semi-predictable such as cognitive test scores, depression status/score, and patient/caregiver desires; and those that are not predictable such as structural MRI findings, safety/diagnostic labs, ECG findings, and the presence of amyloid or tau on CSF or PET imaging. Treating clinicians who are very familiar with their patients and caregivers can increase their awareness of screen failure reasons, rendering those that are semi-predictable as more predictable, and thereby decreasing screen failure rates in clinical trials; whereas sites who rely on advertising for subject recruitment may be at a relative disadvantage.

Thus, a thorough and vigilant approach to predictable causes of screen failures, which includes an examination of general medical history and demographics, specific history and severity of memory impairment, presence of reliable study partner (a caregiver) and subject's willingness to take part in the clinical trials, can dramatically reduce screen failure rate by as much as 10-20%. It is much more difficult to reduce screen failure rates caused by non-predictable factors such as amyloid level on CSF, amyloid-PET or safety brain MRI indicating ARIA.

However, one promising technique that may help to render these factors comparatively more predictable includes utilising statistical tools that help predict the presence of amyloid/tau or even the eventual diagnostic conversion to AD. Typical techniques involve using multiple regression analytic techniques to predict the presence or absence of beta amyloid or tau on imaging or in CSF, based on scores from earlier-obtained, less expensive and easier to acquire screening measures such as demographics, cognitive test scores, genetic status, clinical signs/symptoms and even structural MRI findings.

Another more elaborate and proprietary method has recently been proposed by researchers at the Wisconsin Alzheimer's Disease Research Center and the University of Wisconsin's Geriatric Research Education and Clinical Center. This methodology was designed to minimise the cost of AD trials without compromising statistical power by utilising an adaptive design for data acquisition that exploits harmonic analysis of a band-limited signal on a graph whose node corresponds to participants with the goal of fully recovering a multivariate signal on the nodes, given the full set of lower-cost features and a partial set of more expensive measurements⁶. Specifically, this is accomplished by utilising an adaptive query strategy derived from probing the properties of the graph in the frequency space, a method that can reliably recover the true signal from the participants with only partial observations, directly resulting in accurate predictions about group inclusion (such as those with amyloid PET) as well as offering substantial financial savings. Analytic techniques such as this offer the opportunity to predict which subjects will eventually qualify for study participation in an adaptive manner, with each additional piece of screening information adding to the overall

success of final prediction based on biomarker status that is inherently not predictable without such valuable techniques. These types of analytic methods, along with an increased familiarity of patient clinical status and the use of the aforementioned hierarchical approach to screening, should help to minimise screen failure rates in AD trials, resulting in a corresponding improvement in overall recruitment rates in these notoriously difficult-to-enrol trials.

References

1. Schneider LS, Olin JT, Lyness SA, et al. Eligibility of Alzheimer's disease clinic patients for clinical trials. *Journal of the American Geriatrics Society*. 1997; 45(8):923-928.
2. Treves TA, Verchovsky R, Klimovitsky S, et al. Recruitment rate to drug trials for dementia of the Alzheimer type. *Alzheimer Disease and Associated Disorders*. 2000; 14(4):209-211.
3. Seltzer B, et al. for the Donepezil "402" Study Group. Efficacy of Donepezil in Early-Stage Alzheimer Disease A Randomized Placebo-Controlled Trial. *Arch Neurol*. 2004;61:1852-1856
4. Salloway S et al. Two Phase 3 Trials of Bapineuzumab in Mild-to-Moderate Alzheimer's Disease. *N Engl J Med* 2014;370:322-33.
5. Landau SM et al. Amyloid negativity in patients with clinically diagnosed Alzheimer's disease and MCI. *Neurology* 2016; 86: 1377-1385
6. Johnson, SC., Singh, V. Personal communication Aug 3, 2016.

Tomislav Babic, MD, PhD is Vice President of Medical and Scientific Affairs/Neuroscience Franchise at Worldwide Clinical Trials Inc. Dr Babic is a board-certified neurologist and clinical pharmacologist, with particular interest in drug development for Alzheimer's disease, Parkinson's, and MS. He is the author of more than 60 peer-reviewed articles and books and has been integral to the development of many approved drugs for PD. His expertise has been widely noted in neurodegenerative disorders in both industry and academia for the past 25 years.

Henry J. Riordan, Ph.D. is Executive Vice President of Medical and Scientific Affairs and Global Lead for Neuroscience at Worldwide Clinical Trials. Dr Riordan has been involved in the assessment, treatment and investigation of various CNS drugs and disorders in both industry and academia for the past 20 years. Dr Riordan specialises in clinical trials methodology and has advanced training in biostatistics, experimental design, neurophysiology, neuroimaging and clinical neuropsychology. He has over 100 publications, including co-authoring two books focusing on innovative CNS clinical trials methodology.
Email: henry.riordan@worldwide.com