

CLINICAL RELEVANCE ON ALZHEIMER'S DISEASE ENDPOINTS

C. SAMPAIO

Laboratório de Farmacologia Clínica e Terapêutica e Instituto de Medicina Molecular, Faculdade de Medicina de Lisboa, Universidade de Lisboa, Portugal. Address for correspondence: Prof. Doutor Cristina Sampaio, Laboratório de Farmacologia Clínica e Terapêutica, Faculdade de Medicina de Lisboa, Edifício do Hospital de Santa Maria, Piso 3, elevador 11, 1600 Lisboa, Portugal, Tel: +351217973453, Email: crissampaio@mail.telepac.pt

Introduction

Disease is an abstract concept and Alzheimer's is not different in this particular. However each disease determines impairments and disabilities that are concrete and very real constraints to a complete fulfillment of the patients' life. To be able to evaluate those impairments and disabilities in terms of severity, change with time and effect of interventions there is need to develop measurement tools. Alzheimer's disease encompasses many domains of brain function, namely: cognition, mood, behavior, motorcity and, as a consequence of the previous, functional autonomy. Given the complexity of Alzheimer's phenotype it is rather unlikely that a unique measurement tool is able to capture adequately all the disease domains in its different stages: pre-dementia, mild, moderate and severe dementia. The solution is to use more than a measurement and to attempt an integrated interpretation of the results. Once an actual set of measurements is obtained there is need to assign them a meaning in terms of relevance for the patient life. This assignment of meaning is generically understood as an evaluation of clinical relevance. In a clinical trial setting it is strongly advisable that the criteria that defined clinical relevance are established apriori. This is a procedure linked but it is not synonymous of proposing a difference to power the study. Power calculation has other determinants and an overpowered trial can produce results that are clinically irrelevant.

In this paper we revise the concepts beyond the process of attributing clinical relevance to a result of a clinical study and we give an overview of the data currently available that has used to establish cut-off values. The paper is structured as a narrative review to stimulate the discussion around the methods to establish clinical relevance, yet it was based on a systematic literature search of PubMed concerning Minimal Clinical Important Change (MCIC) and Minimal Clinical Important Difference (MCID).

Concepts: Minimal clinical important change (MCIC) and others

Minimal Clinical Important Change (MCIC) also called Minimal Clinical Important Difference (MCID) for any given measurement is the smallest difference between two assessments that have a perceived impact in disability and/or handicap. It has been also defined as the smallest treatment effect that would result in a change in patient management,

given its side effects, costs and inconveniences (1). The estimation of how big this change must be is a complex matter; it is noted that physician and patient perspectives might be different in this regard. The patient perspective takes precedence but it cannot stand alone given that it is highly difficult to have the insight about what is attributable to main disease and what is a consequence of co-morbidities. In addition given the nature of Alzheimer Disease the caregiver view must be brought to the picture and therefore both should be considered. Established MCICs are overall rare but those established experimentally as opposed to ones established by consensus methods are even more seldom. MCIC has been particularly well studied for pain assessment through VAS (2, 3, 4); MCIC was also determined for Parkinson s disease assessment with the UPDRS [5], for Stroke in relation with the Barthel Index (6), for health status rating scales for asthma and COPD (7) and for a few other conditions but not many, certainly not for Alzheimer disease's outcomes measures. However an exercise in that direction has been made using the tacrine trials which will be comment below (8).

MCIC might have an important inter-individual variability and it might be dependent of the initial scores; in this case MCIC will be lower in mild disease and higher in severe disease. In any event confidence intervals can be established experimentally and might be the most appropriate method to deal with the problem of having probabilistically uncertain cut-off scores for MCIC once these are determine (7). In addition one should not forget the MCIC is scale specific. It cannot be directly transferred from a scale to another.

MCIC can be situated in 2 relevant but different settings: individual and group. The individual level is critical to the definition of responder which is anchored on the change from baseline while the group MCIC is more relevant to interpret changes to a control group at endpoint. In the case of neurodegenerative disorders like Alzheimer Disease the difference from baseline need to take in account the progressive nature of the disease which makes baseline a slippery anchor that needs to be evenhanded by whatever process of discounting.

There are several different methodologies to establish MCIC empirically which have been reviewed recently (9). We will not describe those here but they are substantially different from the exclusive statistically approach proposed by Cohen (10). In the Cohen method differences between 2 groups are standardized as an effect size (mean change of the score divided by the pooled standard deviation); effect sizes can then be

mathematically categorized as irrelevant if < 0.2 , small if ≥ 0.2 and < 0.5 , medium if ≥ 0.5 and < 0.8 and large if ≥ 0.8 . Cohen's categories are independent from clinical parameters consequently they may be very inadequate when applied to a particular clinical situation.

Alzheimer disease outcomes and clinical relevance

Alzheimer Disease impairments and disabilities have been mostly addressed through rating scales either multidomain or domain specific, in which case several scales should be used simultaneously to cover the different aspects of phenomenology of the disease. Global measurements as the CIBIC+, CDR and more recently CDR-SB are of great value. In addition, delays in the occurrence of important natural history milestones are also accepted as relevant endpoints.

The issue of establishing MCIC is more relevant when the measurement is continuous. When the measurement is categorical the MCIC should be at least 1 category. This principle has not been systematically followed in Alzheimer disease trials where categorical global scores have been parameterized and processed to produce means and SD. In this case there are mean group differences which are well below one (1) what makes interpretation difficult once the anchors for the categories are lost. When the endpoint is measured as a survival function the MCIC relates to size of the delay achieved and therefore it is expressed in a time unit.

Currently, the field of trials to test drugs to prove a symptomatic effect is relatively well established yet there are still many open discussions regarding the clinical relevance of their results. For example, one can view the controversy around the recent NICE guidelines (11) that withhold reimbursement for ACHI and memantine from a sizeable fraction of AD patients in England, as politically charged but it is certainly relevant for the discussion of MCIC in Alzheimer disease.

At the moment there are no known results from the ongoing trials aiming to demonstrate a disease modifying effect which are generically set to evaluate a difference in the rate of progression between the treated and the placebo group. The majority of these trials were powered to show a difference of 25 to 30% in that rate of progression. For those that chose ADAS-Cog as one of the primary endpoints, an absolute mean difference to placebo of about 3 points is to be expected at the end of 18 months of follow-up. Will this be considered above MCIC? In the absence of empirically data to inform the judgment a heated debate can be expected.

Below, we give a quick synthesise of data so far available for the AD symptomatic treatments that can inform the definition of a MCIC. The most of the data available is on Cohen effect sizes.

Rockwood K (12) made the exercise of calculating Cohen Effect sizes for the Acetylcholinesterase Inhibitors (ACHI) trials. The median Cohen's d effect sizes (ES) using ITT samples with LOCF for the ADAS-Cog were: low dose of a ChEI (n = 8 studies) median ES = 0.15, range = 0.03–0.22; medium dose (n = 13) median ES = 0.23, range = 0.12–0.29; high dose (n = 9)

median ES = 0.28, range = 0.01–0.31. Global clinical scales produced similar estimates of ES (for example, high dose ChEI, ITT/LOCF median Cohen's d = 0.29, range = 0.20–0.47). Smith et al (13) calculated the Cohen effect sizes for the Memantine trials. In the moderate-severe AD patient population, cognition, as measured by the Severe Impairment Battery, had an effect size of 0.32 and 0.49.

Attempts to establish MCIC for AD endpoints are very scant. Burbach et al (8) evaluated the now historical tacrine trials. His paper is relevant in the sense that they establish the MCIC for the MMSE using the survey-consensus method, in Canada. According with their data the mean MCIC of the MMSE was 3.72 points (95% CI 3.50–3.95). Both the median and mode values of the survey MMSE MCIC were 3.0 points.

Conclusions

Clinical relevance will always be a judgmental matter yet it can be informed by empirically generated data like MCIC or MCID. In the field of Alzheimer Disease this aspect of clinical research has been neglected. We do not support the view defend by Rockwood (12) that Cohen effect sizes give a better insight of the impact of a given trial. Effect sizes take only in account the heterogeneity of the population expressed by the variance. It is strictly a statistic parameter while clinical relevance pertains to objective and subjective perceptions of benefit. We then recommend a planned investment in studies aiming to establish MCIC for the main Alzheimer endpoints in the different settings that are nowadays relevant, namely for long-term trials.

References

1. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407-15.
2. Todd KH, Funk KG, Funk JP, Bonacci R. Clinical significance of reported changes in pain severity. *Ann Emerg Med* 1996;27:485-489.
3. Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. *Ann Emerg Med* 2001;38:633-638.
4. Bird SB, Dickson EW. Clinically significant changes in pain along the visual analog scale. *Ann Emerg Med* 2001;38:639-643.
5. Schrag A, Sampaio C, Counsell N, Poewe W. Minimal clinically important change on the unified Parkinson's disease rating scale. *Mov Disord*. 2006 Aug;21(8):1200-7.
6. Hsieh YW, Wang CH, Wu SC, Chen PC, Sheu CF, Hsieh CL. Establishing the Minimal Clinically Important Difference of the Barthel Index in Stroke Patients. *Neurorehabil Neural Repair*. 2007 Mar 9; [Epub ahead of print]
7. Jones PW. Interpreting thresholds for a clinically significant change in health status in asthma and COPD. *Eur Respir J* 2002; 19: 398-404.
8. Burbach D, Molnar FJ, St. John P, Man-Son-Hing M. Key Methodological Features of Randomized Controlled Trials of Alzheimer's Disease Therapy Minimal Clinically Important Difference, Sample Size and Trial Duration. *Dement Geriatr Cogn Disord* 1999;10: 534-540.
9. Wells G, Beaton D, Shea B et al. Minimal Clinically Important Differences: Review of Methods. *J Rheumatol* 2001;28:406-12.
10. Cohen J. *Statistical power analysis for the behavioural sciences*. New York: academic press, 1977.
11. NICE guideline. Donepezil, galantamine, rivastigmine (review) and memantine for the treatment of Alzheimer's disease. <http://guidance.nice.org.uk/TA111> (consulted on March 1, 2007)
12. Rockwood K. Size of the treatment effect on cognition of cholinesterase inhibition in Alzheimer's disease. *J. Neurol. Neurosurg. Psychiatry* 2004; 75:677-685.
13. Smith M, Wells J, Borrie M. Treatment effect size of memantine therapy in Alzheimer disease and vascular dementia. *Alzheimer Dis Assoc Disord*. 2006 Jul-Sep;20(3):133-7.