

REVIEW

Measuring clinical progression in MCI and pre-MCI populations: enrichment and optimizing clinical outcomes over time

Suzanne B Hendrix*

Abstract

Recent biomarker research has improved the identification of individuals with very early stages of Alzheimer's disease (AD) and has demonstrated that biomarkers are sensitive for measuring progression in the pre-dementia or mild cognitive impairment (MCI) stage and even pre-symptomatic or pre-MCI stage of AD. Because there are no validated biomarkers in AD, it is important to seek out clinical outcomes that are also sensitive for measuring progression in these very early stages of disease. Clinical outcomes are more subjective and more affected by measurement error than biomarkers but represent the core aspects of the disease and are critical for validation of biomarkers and for evaluation of clinical relevance. Identification of individuals with pre-MCI stages of AD will need to continue to rely on biomarkers, but the identification of individuals with MCI who will progress to AD can be achieved with biomarkers or clinical criteria. Although standard clinical outcomes have been shown to be less sensitive to progression than biomarker outcomes in MCI and pre-MCI populations, non-standard scoring has improved the performance of the Alzheimer's Disease Assessment Scale cognitive subscale, making it more sensitive to progression. Neuropsychological cognitive testing items are optimal for measuring progression in pre-MCI populations, and current research is exploring the best ways to combine these items into a composite cognitive score with maximum responsiveness. In an MCI stage, cognitive, functional, and global items all change, and the best single composite score for measuring progression may involve all of these aspects of the disease. The best chance of success in demonstrating treatment effects in clinical trials will be achieved in a well-defined pre-MCI or MCI population and with an outcome that tracks well with clinical progression over time and with time. A partial least squares model can be used to identify these optimal weighted combinations.

Introduction

The development of Alzheimer's disease (AD) therapies with putative disease-modifying mechanisms has led to the strategy of testing these therapies in early stages of the disease in order to give them the best chance of affecting the disease before it is fully established. As a result, recent research has focused on exploring methods for potential enrichment of a subject population on the basis of identifying predictors of future decline or 'conversion' to mild cognitive impairment (MCI) or AD and selecting optimal outcome measures for measuring progression in these populations. Although most research focuses on one of these approaches, combining enrich-

ment approaches with optimized outcome measures is likely to further increase power in clinical trials.

The aims of this paper are to describe the current approaches for enriching patient populations in MCI and pre-MCI populations, describe the current approaches to measuring progression of patients over time, and propose an approach to developing a clinical composite score with improved responsiveness to clinical progression. In addition, some guidelines and cautions are suggested for evaluating the ability of an outcome to measure disease progression. As we consider clinical trial design, the available research helps us make decisions of two different types: how do we enrich the subject population being selected for the study, and what is the best way to measure the progression of subjects over time? Of course, these two decisions are inextricably linked; the best method for measuring progression of subjects over time will depend on the subjects who are being studied. Both biomarkers and clinical outcomes have been considered

*Correspondence: shendrix@pentaracorp.com
Pentara Corporation, 2180 Claybourne Avenue, Salt Lake City, UT 84109, USA

for each of these two applications, and generally, decisions will be made in a way that minimizes the sample size requirement for a clinical study, but it is helpful to evaluate the assumptions behind this criterion for evaluation.

A detailed description of the results of the research that drives these clinical trial design decisions is beyond the scope of this review. (See [1-3] for an overview of the research findings.) Instead, the focus will be on describing the methods for identifying populations and the methods for developing new clinical composites for measuring progression in MCI and pre-MCI populations in support of clinical trial design decisions. Also, clinical measures rather than biomarker measures will be emphasized (see [4] for a detailed discussion of the use of biomarkers in AD drug development), and some challenges in interpreting the literature in this area will be addressed.

Enrichment

Biomarkers are often used for subject selection in clinical trials, particularly in early disease. Recent biomarker research has supported additional criteria for identifying individuals who have MCI and are most likely to progress to AD. These criteria include molecular biomarkers such as cerebrospinal fluid (CSF) A β -42 of below approximately 192 pg/mL, higher values of the CSF tau/A β -42 ratio, and reduced glucose metabolism demonstrated with 2-[18F]-fluoro-2-deoxy-D-glucose-positron emission tomography (FDG-PET) imaging [1,3,5]. Biomarkers that detect A β deposition, such as CSF A β -42, the CSF tau/A β -42 ratio, and [11C] Pittsburgh compound B (PiB) PET, are useful for identifying healthy subjects who are likely to progress to MCI [6]. FDG-PET may also be helpful in this early stage [7].

Selection of a population for study in a clinical trial is an important process that is usually intended to target subjects who will convert to MCI or AD within a certain amount of time and with a degree of certainty. Although subjects who will convert within a certain time frame cannot be prospectively selected with certainty, the retrospective separation of converters and non-converters allows a comparison of decline rates between those who are close to a diagnosis of MCI or AD and those who are not. Another approach that can be used prospectively or retrospectively is to separate subjects who are in a pre-MCI stage into groups on the basis of a DNA marker such as apolipoprotein E gene e4 allele (*APOE-e4*) or presenilin 1 (*PSI*) gene carrier status. Each of these approaches can be used in conjunction with optimizing a clinical outcome to be sensitive to decline over time. If AD is a single entity regardless of whether its occurrence is sporadic or genetic, then the combinations of items that are most sensitive to change will be similar with each of these different approaches.

The pre-MCI stage of AD is characterized by changes in biomarkers such as volumetric magnetic resonance imaging (MRI), CSF tau, and CSF A β -42 levels and functional MRI. Biomarkers are more suited than clinical markers to identifying individuals with pre-MCI AD. This is not because of the complete absence of clinical changes prior to a clinical diagnosis of MCI but because of the highly variable nature of the neuropsychological changes that are seen in this very early population. This large variability could be partly overcome by following subjects longitudinally and observing changes within a subject, but biomarkers naturally lend themselves to use in the selection process because of their objective nature. Also, an enrichment biomarker does not require the same validation that would be required for a biomarker to be used as an outcome assessment (Table 1).

Although biomarkers are better than clinical outcomes for identifying individuals in a pre-MCI stage, several authors [8-10] have shown that cognitive outcomes are able to compete with biomarker outcomes in identifying individuals in an MCI stage. In fact, Llano and colleagues [8] developed a weighted version of the Alzheimer's Disease Assessment Scale cognitive subscale (ADAS-cog) that performed better than any single MRI measure in predicting progression from MCI to AD in the Alzheimer's Disease Neuroimaging Initiative (ADNI) data set. In a pre-MCI population, decline on neuropsychological tests or a score that is lower than a threshold may also be predictive of subjects who are likely to progress to MCI.

It is important to note that the goal of enrichment is to include individuals who eventually will progress to AD and to exclude those who will not. Over-enriching a population by selecting only subjects who are nearly certain to progress to AD at a rapid rate may result in what looks like a more powerful clinical study, but the results are not likely to be generalizable to a broader population and we may be excluding the subjects whose progression is slow enough that we have time to intervene.

Biomarkers as outcomes

In the setting of clinical trial design, the anticipated sample size necessary for an outcome to detect a treatment effect is critically important and has been used as a standard of comparison for longitudinal outcomes. The standard estimate that has been used for this comparison is the sample size per arm required to detect a 25% reduction in atrophy/decline with 80% power and 5% significance [3]. This criterion reflects the sensitivity to decline or external responsiveness of a clinical outcome, which is the ability of a clinical outcome to change with time and disease progression. External responsiveness takes a pre-eminent role in evaluating clinical trial outcomes because the nature of AD, even in its very early stages, is degenerative, and the aspects of the disease

Table 1. Options for enrichment and measuring progression in pre-mild cognitive impairment and mild cognitive impairment

		Enrichment	
		Biomarkers/ DNA markers	Clinical measures
Measuring progression	Biomarkers	Biomarker not validated as outcome	Clinical measures are too variable, and biomarker is not validated as outcome.
	Clinical outcomes	Best option	Measurement error leads to potential issues with regression toward the mean.

which degrade over time offer the most promise in terms of outcomes that reflect the disease process.

Although CSF biomarkers and amyloid imaging seem to be the best at classifying or selecting individuals, volumetric MRI outcomes seem to be sensitive biomarker measures of disease progression in MCI and pre-MCI populations. Whole-brain volume and hippocampal volume have been shown to be more sensitive to longitudinal disease progression than cognitive measures – ADAS-cog and mini-mental status examination (MMSE) – in an MCI population [11-13]. Some of the MRI regions that show sensitivity to decline in an MCI population are also sensitive to decline in a pre-MCI population.

Many studies in MCI demonstrate that a smaller sample size is required with a biomarker such as volumetric MRI as the primary outcome than for studies with a primary clinical outcome [3]; however, this does not imply that clinical outcomes are not important in these early stages. In fact, it may be harder to change an outcome such as volumetric MRI than to change a clinical outcome, particularly if a treatment has both a disease-modifying and a symptomatic effect. The ability of an outcome to change with a treatment effect is referred to as internal responsiveness or as sensitivity to treatment effects. If volumetric MRI has less internal responsiveness than a clinical outcome, factoring this into the sample size calculations would reduce the apparent advantage that volumetric MRI has over clinical outcomes. Also, relying exclusively on a biomarker outcome has the risk that a treatment effect on a biomarker may not translate into a treatment effect on a clinical outcome.

For both of these reasons, it is advantageous to measure standard clinical outcomes in studies that have a biomarker as a primary outcome, even though the studies may not be powered to show significance on a clinical effect. Additionally, studies with both biomarker and clinical outcomes will facilitate future validation of biomarker outcomes and will provide data to support development or testing of future composite clinical outcomes combining items from standard instruments.

Evaluation of clinical progression outcomes

No standard clinical outcomes are currently established in MCI and pre-MCI populations. Any outcomes proposed for use will need to be validated in the relevant population in order to be used as a primary outcome in a pivotal study for regulatory submission. The validation process typically includes demonstrating reliability and validity. In addition, the responsiveness of the scale, both external and internal, should be assessed [14,15] (The article by Coley and colleagues [15] interprets internal and external responsiveness differently.) Because this field is rich in reliable and validated neuropsychological tests (including cognitive outcomes that measure many different cognitive domains and outcomes that measure function and global changes), the focus should be on improving responsiveness as the primary challenge in measuring progression in MCI and pre-MCI populations. This focus does not ignore the validation requirement for a new clinical composite outcome, but merely emphasizes responsiveness as the area of greatest challenge with a very slowly declining population.

Outcomes that have been proposed and used in these very early populations include single neuropsychiatric tests originally used to measure deviations from normal cognition, such as the Free and Cued Selective Reminding test, and outcome measures that are commonly used in mild-to-moderate AD, such as the ADAS-cog and Clinical Dementia Rating sum of boxes (CDR-sb). These outcome measures have good reliability and validity [16,17] within the populations for whom they were developed but may not have optimal responsiveness in MCI and pre-MCI since the outcomes were not developed specifically for the longitudinal monitoring of cognitive changes in a slowly progressive, mildly impaired population. Even clinical outcomes with changes that are highly predictive of progression to AD or MCI may not be the most responsive outcomes longitudinally in populations that we are able to define prospectively.

Internal responsiveness is inherently difficult to estimate since current therapies are thought to be

symptomatic and therapies in development in these early stages are hoped to be disease-modifying. The cognitive items that respond to a disease-modifying treatment may not be the same items that respond to a symptomatic treatment. It seems reasonable to assume that a disease-modifying treatment would be expected to slow all aspects of the disease by the same percentage, but this assumption may not hold if some outcomes are more reversible than others and perhaps more easily slowed. Biomarkers may not be as sensitive to slowing as clinical outcomes, even for a purely disease-modifying treatment, if slowing clinical outcomes also results in indirect clinical benefit because of improved subject or caregiver outlook. Because these issues are complex, consideration of different scenarios of internal and external responsiveness is important, and using a measure of sensitivity to decline, such as the mean to standard deviation ratio (MSDR) [18] or its reciprocal (the coefficient of variation), allows the estimation of sample size for several different scenarios. Table 2 provides a simple sample size table in order to facilitate translation between these different ways of comparing sensitivity to decline. (The sample sizes reported in the literature with the scenario of 25% effect can be looked up on the 25% row to get the estimated MSDR from the column header.)

In the same way that items may not be equally responsive to a treatment effect, two different subject populations may not be equally responsive to a treatment effect. Comparing the power/sample size between two populations defined by different criteria for enrichment assumes that the treatment effect size will be the same within the two enriched groups. This assumption is impossible to test but seems to be reasonable if the purpose of enrichment is to separate out MCI converters from MCI non-converters or pre-AD MCI from other MCI. If the enrichment is being used instead to select a group of fast decliners, it seems unlikely that a disease-modifying treatment effect would be as large for faster decliners as it would be for slower decliners. In this case, any estimated improvements in power/sample size may be misleading since the reduced treatment effect size may counteract those improvements.

Developing a responsive outcome with modeling

Evaluating external responsiveness of a clinical scale requires a 'gold standard' of health status. Using future decline on a standard clinical outcome, such as the CDR-sb or ADAS-cog, as the gold standard or a future 'conversion' endpoint requires a retrospective approach that may not be as applicable to a population enrolled in a clinical trial. A principal components analysis on the change scores uses the overall direction of the clinical changes as the gold standard. Using the MSDR of different composite scores as a criterion for selection in

an exhaustive search approach equates to using time as the gold standard of health status [19]. A modeling approach using an ordinary least squares (OLS) model with time as the outcome variable and selecting items that are predictive of time is a direct way to achieve what the exhaustive search method seeks: to create a composite score that optimizes responsiveness over time. A partial least squares (PLS) approach that uses time as the dependent variable and the item scores as the independent variables combines the best of both of these approaches by identifying a weighted combination of items which is associated with time and decline in the clinical scores.

To use time or overall clinical decline as the gold standard, the population selected for inclusion in the study should be well defined as a group that has AD. For the MCI and pre-MCI populations, the subsequent diagnosis can be used retrospectively to define the population that is then compared in terms of external responsiveness (that is, sensitivity to decline over time). Although internal responsiveness is also important, a test that is not sensitive to decline over time is not likely to be responsive to treatment effects, particularly treatment effects that slow the disease progression. Outcomes that are currently used in these disease stages could be compared with new outcomes in order to see whether the new outcome provides improved sensitivity to decline.

The ADAS-cog is well established as an outcome measure in mild-to-moderate AD but clearly has several items that are not expected to change in early disease [20]. Including these items in the ADAS-cog can hurt its performance in terms of external responsiveness over time. Different weighting of ADAS-cog items in order to minimize the impact of these less sensitive items or even eliminate these items from the scale results in a cognitive composite with improved sensitivity in measuring progression over time in MCI subjects [19]. A combination score that allows inclusion of items from neuropsychological testing, traditional cognitive tests such as the ADAS-cog and MMSE, functional assessments such as Alzheimer's Disease Cooperative Study-activities of daily living (ADCS-ADL) and Disability Assessment for Dementia (DAD), and global assessments such as the CDR-sb, CIBIC+ (Clinician Interview-Based Impression of Change, plus career interview), or ADCS-Clinical Global Impression of Change (ADCS-CIGIC) would likely increase the performance even further. Although it seems unusual to combine items that measure different domains of the disease, this approach reflects the belief that AD, prior to dementia, is a single entity that can be measured with a single combination score. Use of a global score such as the CDR-sb as a single primary outcome also reflects that belief, but this global score may be

Table 2. Sample sizes for different treatment effect sizes and different scale sensitivities to decline

Sample size per group	Mean-to-standard deviation ratio of change over time (coefficient of variation)								
	0.2 (5)	0.3 (3.3)	0.4 (2.5)	0.5 (2)	0.6 (1.7)	0.7 (1.4)	0.8 (1.3)	0.9 (1.1)	1 (1)
Percentage effect relative to placebo									
20	9,812	4,361	2,453	1,570	1,091	802	614	485	393
25	6,280	2,791	1,570	1,005	698	513	393	311	253
30	4,361	1,939	1,091	698	485	357	274	217	176
35	3,204	1,424	802	513	357	263	202	160	130
40	2,453	1,091	614	393	274	202	155	123	100
45	1,939	862	485	311	217	160	123	97	79
50	1,570	698	393	253	176	130	100	79	64
55	1,298	577	325	209	146	107	83	66	53
60	1,091	485	274	176	123	90	70	55	45
65	929	413	234	150	105	77	60	47	39
70	802	357	202	130	90	67	52	41	34
75	698	311	176	113	79	58	45	36	29
80	614	274	155	100	70	52	40	32	26
85	544	243	137	88	62	46	35	28	23
90	485	217	123	79	55	41	32	25	21
95	435	195	110	71	50	37	29	23	19
100	393	176	100	64	45	34	26	21	17

enhanced by the addition of cognitive or functional items.

Statistical modeling can be used to find the best weighted combination of items for measuring progression in a given population. Owing to the highly variable nature of MCI populations and pre-MCI populations, any combination identified by using the methods above would also need to be validated across multiple studies and different populations. Cross-validation using re-sampling from pooled data of multiple studies in different patient populations is preferable to using a single study to validate another single study when there are between-study differences. A single outcome that measures the strongest dimension of disease-related decline in a very early population would be extremely valuable, particularly in a proof-of-concept study in which the primary goal is to determine whether a treatment has promise. The addition of a clinical outcome for decision making rather than reliance on a biomarker outcome alone reduces the risk of moving into pivotal studies.

Challenges in enrichment and outcome assessment

Several challenges should be kept in mind when deciding if and how to enrich a patient population for inclusion in a clinical trial and when selecting the best tool for measuring change in that patient population. Many of these challenges can be easily addressed once they are understood.

As discussed above, clinical assessments can be used to enrich a patient population. However, owing to regression toward the mean, using a clinical outcome to identify subjects for a study and a related clinical outcome to follow those same subjects over time is not ideal. The subjects who perform poorly at entry into the study because of measurement error are more likely to have less penalty due to measurement error at the next visit, resulting in less decline than would be expected. This effect can be reduced either by following subjects for a long enough time period (2 years or more) after enrollment or by using a clinical outcome that is not closely related to the outcomes used for enrichment of the subject population.

Enrichment is intended to either maximize the chance of progression to the next stage of AD which is a dichotomous outcome or maximize the degree of progression which is a continuous outcome. The difference in power between these two approaches comes down to the question of whether 'conversion' to MCI or AD is really a dichotomy or a progression to a somewhat arbitrary threshold. In a population that includes subjects who will never progress to AD and others who will progress to AD (such as a healthy population), it could be argued that conversion may be a more appropriate outcome. But if we have enriched such that most or all subjects in our study are expected to shift closer to conversion within the time of the study and eventually

will progress to the next stage of AD, such as in an MCI population, then it is more powerful to measure decline as a continuous outcome [21].

If an outcome has been optimized on the basis of a particular study population, then it is important to assess whether the population being selected for the study is similar to the study population used to optimize the outcome. If the populations are not similar but the study population is predictably different from the original population, one could do additional modeling to account for these differences. For instance, if the optimal combination differed depending on time to progression to the next stage of dementia, the expected distribution of time prior to progression could be used to weight the combinations in order to get a reasonable estimate of the overall decline rate that would be expected in the study population and in order to ensure that the outcome measure is optimized for the population being enrolled.

When the decline of an outcome over time is considered, it is important to consider the normal aging decline over time in healthy subjects. Correcting for normal aging may be particularly important when the healthy group declines over time since a disease-modifying treatment effect is not likely to be able to slow normal aging effects even when it has a slowing effect on disease-specific decline. Correcting for normal aging may be less important when the healthy group has learning effects over time, since the estimates of decline over time will be conservative in this case. If a healthy control group is included in the study, then a correction for normal aging can be done on a group level or on an individual level on the basis of the specific age of the individual and a model fit to the healthy control group. In general, both corrected and uncorrected analyses should be considered when possible.

Conclusions

In a pre-MCI population, retrospective enrichment based on comparing those who progress to MCI with those who do not or on comparing mutation carriers with non-carriers offers the best setting in which to optimize a clinical outcome for measuring progression based on external responsiveness to changes over time. The combination of identifying a population of subjects who could be prospectively enrolled in a clinical study – such as MCI subjects, prodromal MCI subjects, or mutation carriers – and then optimizing a clinical outcome in that population results in a composite score that has the best chance for maximal power in a clinical study with these specific MCI populations. In both of these approaches, cross-validation is important and can be performed across different data sets if the populations are similar enough or with split-sample validation methods applied to pooled samples when study populations differ.

Quantitative outcomes are likely to be more powerful than dichotomous endpoints since only large changes are captured with dichotomous endpoints and more subtle changes can be seen with quantitative outcomes. Although it could be argued that we may observe statistically significant differences that are not clinically relevant, it is important to remember that the declines that are observed in an MCI population and particularly in a pre-MCI population are subtle enough that the average progression in this stage of the disease may not be considered clinically relevant using historical standards.

A PLS approach using time as the dependent variable and the item scores as the independent variables essentially uses time and clinical decline as the 'gold standard,' combining the best attributes of a principal components approach with the best attributes of an exhaustive search or OLS approach based on time. Using the MSDR to identify the best composite score results in selection of an outcome measure that tracks most sensitively with progression because the MSDR measures the external responsiveness to time. This approach is based on the placebo group decline and assumes a constant percentage reduction in the active group. So this approach is more appropriate for a disease-modifying treatment that may be expected to impact all clinical disease progression similarly than for a symptomatic treatment that is likely to have a larger effect on some symptoms than on others.

Composite cognitive scales that combine items from neuropsychological tests offer improved measurement of decline in a pre-MCI population. In an MCI stage, a composite that considers cognitive, functional, and global items is likely to give the best chance for optimal measurement of decline, reflects a global approach to the disease, and would be particularly useful for proof-of-concept studies. A composite that is restricted to cognitive items in an MCI stage would offer improvement over standard clinical outcomes and would offer the simplicity of measuring a single domain of progression. Either of these two approaches could be used in parallel with an appropriate biomarker outcome such as volumetric MRI for a treatment that is expected to slow disease progression.

The combination of enrichment of the study population and optimization of sensitivity to decline by using a weighted composite score gives us the best chance to improve the efficiency of a clinical trial in a pre-MCI or MCI population. This improved efficiency allows us to perform shorter, smaller studies than would otherwise be required. In addition, this approach could give us more confidence in a positive or negative result or allow us to get a more accurate estimate of a treatment effect in an inconclusive proof-of-concept study.

Abbreviations

AD, Alzheimer's disease; ADAS-cog, Alzheimer's Disease Assessment Scale cognitive subscale; ADCS, Alzheimer's Disease Cooperative Study; CDR-sb, Clinical Dementia Rating, sum of boxes; CSF, cerebral spinal fluid; FDG-PET, 2-[18F]-fluoro-2-deoxy-D-glucose-positron emission tomography; MCI, mild cognitive impairment; MMSE, mini-mental status examination; MRI, magnetic resonance imaging; MSDR, mean to standard deviation ratio; OLS, ordinary least squares; PET, positron emission tomography; PLS, partial least squares.

Competing interests

SBH is president and owner of Pentara Corporation (Salt Lake City, UT, USA), a consulting firm that consults with several pharmaceutical companies and non-profit groups that are conducting clinical trials in AD.

Published: 13 July 2012

References

- Brooks LG, Loewenstein DA: **Assessing the progression of mild cognitive impairment to Alzheimer's disease: current trends and future directions.** *Alzheimers Res Ther* 2010, **2**:28.
- Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, Iwatsubo T, Jack CR Jr., Kaye J, Montine TJ, Park DC, Reiman EM, Rowe CC, Siemers E, Stern Y, Yaffe K, Carrillo MC, Thies B, Morrison-Bogorad M, Wagster MV, Phelps CH: **Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.** *Alzheimers Dement* 2011, **7**:280-292.
- Weiner MW, Veitch DP, Aisen PS, Beckett LA, Cairns NJ, Green RC, Harvey D, Jack CR, Jagust W, Liu E, Morris JC, Petersen RC, Saykin AJ, Schmidt ME, Shaw L, Siuciak JA, Soares H, Toga AW, Trojanowski JQ: **The Alzheimer's Disease Neuroimaging Initiative: a review of papers published since its inception.** *Alzheimers Dement* 2012, **8** (1 Suppl):S1-S68.
- Cummings J: **Biomarkers in Alzheimer's disease drug development.** *Alzheimers Dement* 2011, **7**:e13-44.
- Shaw LM, Vanderstichele H, Knapiak-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, Lewczuk P, Dean R, Siemers E, Potter W, Lee VM, Trojanowski JQ: **Alzheimer's Disease Neuroimaging Initiative: Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects.** *Ann Neurol* 2009, **65**:403-413.
- Ho AJ, Hua X, Lee S, Leow AD, Yanovsky I, Gutman B, Dinov ID, Loporé N, Stein JL, Toga AW, Jack CR Jr., Bernstein MA, Reiman EM, Harvey DJ, Kornak J, Schuff N, Alexander GE, Weiner MW, Thompson PM: **Alzheimer's Disease Neuroimaging Initiative: Comparing 3 T and 1.5 T MRI for tracking Alzheimer's disease progression with tensor-based morphometry.** *Hum Brain Mapp* 2010, **31**:499-514.
- Hill D: **Neuroimaging to assess safety and efficacy of AD therapies.** *Expert Opin Investig Drugs* 2010, **19**:23-26.
- Llano DA, Laforet G, Devanarayan V: **Derivation of a new ADAS-cog composite using tree-based multivariate analysis: prediction of conversion from mild cognitive impairment to Alzheimer disease.** *Alzheimer Dis Assoc Disord* 2010, **25**:73-84.
- Fleisher AS, Sun S, Taylor C, Ward CP, Gamst AC, Petersen RC, Jack CR Jr., Aisen PS, Thal LJ: **Volumetric MRI vs clinical predictors of Alzheimer disease in mild cognitive impairment.** *Neurology* 2008, **70**:191-199.
- Jack CR Jr., Petersen RC, Xu YC, O'Brien PC, Smith GE, Ivnik RJ, Boeve BF, Waring SC, Tangalos EG, Kokmen E: **Prediction of AD with MRI-based hippocampal volume in mild cognitive impairment.** *Neurology* 1999, **52**:1397-1407.
- Schuff N, Woerner N, Boreta L, Kornfield T, Shaw LM, Trojanowski JQ, Thompson PM, Jack CR Jr., Weiner MW: **Alzheimer's Disease Neuroimaging Initiative: MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers.** *Brain* 2009, **132** (Pt 4):1067-1077.
- Holland D, Brewer JB, Hagler DJ, Fennema-Notestine C, Dale AM: **Alzheimer's Disease Neuroimaging Initiative: Subregional neuroanatomical change as a biomarker for Alzheimer's disease.** *Proc Natl Acad Sci U S A* 2009, **106**:20954-20959.
- Beckett LA, Harvey DJ, Gamst A, Donohue M, Kornak J, Zhang H, Kuo JH: **Alzheimer's Disease Neuroimaging Initiative: The Alzheimer's Disease Neuroimaging Initiative: annual change in biomarkers and clinical outcomes.** *Alzheimers Dement* 2010, **6**:257-264.
- Husted JA, Cook RJ, Farewell VT, Gladman DD: **Methods for assessing responsiveness: a critical review and recommendations.** *J Clin Epidemiol* 2000, **53**:459-468.
- Coley N, Andrieu S, Jaros M, Weiner M, Cedarbaum J, Vellas B: **Suitability of the Clinical Dementia Rating-Sum of Boxes as a single primary endpoint for Alzheimer's disease trials.** *Alzheimers Dement* 2011, **7**:602-610.e2.
- Grober E, Ocepek-Welikson K, Teresi JA: **The Free and Cued Selective Reminding Test: evidence of psychometric adequacy.** *Psychology Science Quarterly* 2009, **51**:266-282.
- Rosen WG, Mohs RC, Davis KL: **A new rating scale for Alzheimer's disease.** *Am J Psychiatry* 1984, **141**:1356-1364.
- Hendrix SB: **Requiring an amyloid-β1-42 biomarker may improve the efficiency of a study, and simulations may help in planning studies.** *Alzheimers Res Ther* 2011, **3**:10.
- Shankle R, Atri A, Hendrix S: **Improving measurement methodology to detect treatment effect in clinical trials.** *J Nutr Health Aging* 2011, **15** (Suppl 1):s3. Abstract.
- Cano SJ, Posner HB, Moline ML, Hurt SW, Swartz J, Hsu T, Hobart JC: **The ADAS-cog in Alzheimer's disease clinical trials: psychometric evaluation of the sum and its parts.** *J Neurol Neurosurg Psychiatry* 2010, **81**:1363-1368.
- Donohue MC, Gamst AC, Thomas RG, Xu R, Beckett L, Petersen RC, Weiner MW, Aisen P: **Alzheimer's Disease Neuroimaging Initiative: The relative efficiency of time-to-threshold and rate of change in longitudinal data.** *Contemporary Clin Trials* 2011, **32**:685-693.

doi:10.1186/alzrt127

Cite this article as: Hendrix SB: Measuring clinical progression in MCI and pre-MCI populations: enrichment and optimizing clinical outcomes over time. *Alzheimer's Research & Therapy* 2012, **4**:24.