

RESEARCH ARTICLE

Open Access



Using a distribution-based approach and systematic review methods to derive minimum clinically important differences

Jennifer A. Watt^{1,2*}, Areti Angeliki Veroniki^{1,3,4}, Andrea C. Tricco¹ and Sharon E. Straus^{1,2,5}

Abstract

Background: Clinical interpretation of changes measured on a scale is dependent on knowing the minimum clinically important difference (MCID) for that scale: the threshold above which clinicians, patients, and researchers perceive an outcome difference. Until now, approaches to determining MCIDs were based upon individual studies or surveys of experts. However, the comparison of meta-analytic treatment effects to a MCID derived from a distribution of standard deviations (SDs) associated with all trial-specific outcomes in a meta-analysis could improve our clinical understanding of meta-analytic treatment effects.

Methods: We approximated MCIDs using a distribution-based approach that pooled SDs associated with baseline mean or mean change values for two scales (i.e. Mini-Mental State Exam [MMSE] and Alzheimer Disease Assessment Scale – Cognitive Subscale [ADAS-Cog]), as reported in parallel randomized trials (RCTs) that were included in a systematic review of cognitive enhancing medications for dementia (i.e. cholinesterase inhibitors and memantine). We excluded RCTs that did not report baseline or mean change SD values. We derived MCIDs at 0.4 and 0.5 SDs of the pooled SD and compared our derived MCIDs to previously published MCIDs for the MMSE and ADAS-Cog.

Results: We showed that MCIDs derived from a distribution-based approach approximated published MCIDs for the MMSE and ADAS-Cog. For the MMSE (51 RCTs, 12,449 patients), we derived a MCID of 1.6 at 0.4 SDs and 2 at 0.5 SDs using baseline SDs and we derived a MCID of 1.4 at 0.4 SDs and 1.8 at 0.5 SDs using mean change SDs. For the ADAS-Cog (37 RCTs, 10,006 patients), we derived a MCID of 4 at 0.4 SDs and 5 at 0.5 SDs using baseline SDs and we derived a MCID of 2.6 at 0.4 SDs and 3.2 at 0.5 SDs using mean change SDs.

Conclusion: A distribution-based approach using data included in a systematic review approximated known MCIDs. Our approach performed better when we derived MCIDs from baseline as opposed to mean change SDs. This approach could facilitate clinical interpretation of outcome measures reported in RCTs and systematic reviews of interventions. Future research should focus on the generalizability of this method to other clinical scenarios.

Keywords: Meta-analysis, Systematic review, Minimum clinically important difference, Back-transformation, Mini-mental state exam, Alzheimer disease assessment scale – cognitive subscale

* Correspondence: jennifer.watt@utoronto.ca

¹Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital- Unity Health Toronto, 209 Victoria Street, East Building, Room 723, Toronto, Ontario M5B 1W8, Canada

²Division of Geriatric Medicine, Department of Medicine, University of Toronto, 190 Elizabeth Street, R. Fraser Elliott Building, 3-805, Toronto, Ontario M5G 2C4, Canada

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Introduction

In communicating research findings to knowledge users (e.g. patients, caregivers, clinicians), researchers must describe the statistical and clinical significance of their findings, which can be challenging when changes in health status are reported with a clinical scale. For example, although cholinesterase inhibitors and memantine are associated with statistically significant improvements in cognitive function in persons with dementia, the clinical meaningfulness derived from treatment with cholinesterase inhibitors and memantine is unclear [1, 2]. The clinical meaningfulness of changes measured on a scale is dependent on knowing the minimum clinically important difference (MCID) for that scale: the threshold above which clinicians, patients, and researchers perceive an outcome difference [3]. It is challenging for clinicians to discuss the clinical importance of research findings with patients when the MCID for a scale is unknown; shared decision making is inadequate without this information.

There are two main approaches for determining the MCID: anchor-based and distribution-based [4]. An anchor-based approach compares the change in a scale-based outcome measure with that of a patient-reported outcome (e.g. global ratings of change) or other external criterion (e.g. expert opinion, clinical test result) [4–8]. For example, clinical experts agreed that a difference of 1 to 2 points on the Mini-Mental State Exam (MMSE), a test that measures change in cognitive function, was clinically important in a trial comparing the effects of cognitive enhancing medications (donepezil and memantine) to placebo in persons with Alzheimer disease [6]. A distribution-based approach compares the difference in a scale-based outcome measure to a pre-specified threshold value of its uncertainty (e.g. standard error, standard deviation [SD]), which facilitates MCID derivation when direct patient or clinician input is not readily accessible [4, 6]. Cohen proposed that 0.2 SDs represented a small difference and 0.8 SDs represented a large difference [9]. A range of 0.4 to 0.5 SDs is felt to be clinically meaningful and previous work has shown that most MCIDs are within 0.5 SDs [6, 10, 11]. Using a distribution-based approach at a threshold of 0.4 SDs, Howard et al., estimated a MMSE MCID of 1.4 points using SDs for mean change MMSE scores and a MMSE MCID of 1.7 points using SDs for baseline MMSE scores [6]. Although these MCID estimates for the MMSE, which were derived with anchor- and distribution-based approaches, are in agreement, this is not always the case [6, 12]. For example, clinicians' opinions will be shaped by the patients in their practice, outlying outcomes (better or worse than expected) among their patient population, and more recent outcomes experienced by patients [13]. Prospectively comparing a patient-reported outcome to a

scale-based outcome will help to overcome this problem, but estimates are still based on a single sample of patients. Disagreements between clinical experts about appropriate MCIDs and the need to have MCIDs that reflect a wide range of patients are reasons why we need more robust approaches to calculating anchoring bias-free MCIDs [6, 12]. There is no preferred method for establishing the MCID.

Until now, approaches for determining MCIDs were based upon individual studies or surveys of experts [4, 6, 12]. These methods may be appropriate if researchers wish to derive MCIDs for participants of a particular randomized trial (RCT). However, the comparison of meta-analytic treatment effects to a MCID derived from a distribution of standard deviations (SDs) associated with all trial-specific outcomes in a meta-analysis could improve our clinical understanding of meta-analytic treatment effects. Furthermore, the calculation of MCIDs based on a systematic review could enhance clinical decision-making when the MCID for a scale is unknown. We propose a distribution-based approach that approximates MCIDs for continuous outcomes reported in a systematic review of RCTs, which we illustrate with two empiric examples.

Methods

Data set

We used data from a published systematic review and network meta-analysis of the comparative effectiveness and safety of cognitive enhancers (donepezil, galantamine, rivastigmine and memantine) for treating Alzheimer disease [14]. Specifically, we used data on the comparative efficacy of cognitive enhancers for improving the MMSE (56 RCTs) and Alzheimer Disease Assessment Scale – Cognitive Subscale (ADAS-Cog) (53 RCTs) score of persons with Alzheimer disease [14–16]. We included parallel RCTs from each systematic review dataset reporting a baseline mean or mean change value for the MMSE or 11-item version of the ADAS-Cog, SD values for the baseline mean or mean change scale score, and number of participants per study arm reporting this data [16]. We used accepted methods to calculate SDs where study authors reported other measures of uncertainty (i.e. 95% confidence interval or standard error) [17].

Calculating a minimum clinically important difference from pooled standard deviations in a systematic review

We followed these steps to derive MCIDs for MMSE and ADAS-Cog scales:

- a) Derived a pooled SD (SD_{pooled}) from parallel RCTs included in a systematic review reporting the scale of interest, where n_i is the number of participants per study arm, and SD_i is the standard deviation

associated with each mean change or baseline scale score per study arm [18]:

$$SD_{\text{pooled}} = \sqrt{\frac{\sum (n_i - 1) SD_i^2}{\sum (n_i - 1)}}$$

This method for pooling SDs was suggested by Furu-kawa et al. [18]. In a systematic review and pairwise meta-analysis, there is only one treatment comparison and there are two treatment arms. In a systematic review and network meta-analysis, there are two or more treatment comparisons and there could be two or more treatment arms. A pooled SD could be derived across all treatment arms or across each specific treatment arm using this method.

- b) Multiplied SD_{pooled} by an appropriate threshold for SD values to derive a range of plausible values for the MCID [6]. A range between 0.4 and 0.5 SDs is felt to be clinically meaningful and most published MCIDs fall within 0.5 SDs [6, 10, 11].

MCIDs based upon pooled SDs associated with mean change scale scores (i.e. follow-up time point scale score compared to baseline scale score) were also derived using the aforementioned steps; however, SD_i was the SD associated with each mean change scale score per study arm. We derived MCID values at 0.4 and 0.5 SDs to represent the range of clinically meaningful MCIDs. In the primary analysis, we included data from all treatment groups included in the systematic review and network meta-analysis (i.e. donepezil, rivastigmine, galantamine, memantine, and placebo). We performed a sensitivity analysis where SDs estimated from other measures of uncertainty (i.e. 95% confidence interval, standard error) were removed from the pooled SD. In a secondary analysis, we derived MCIDs for each treatment group separately.

Results

Pooled baseline SDs, as described in Table 1, were larger than pooled mean change SDs. MCIDs were unchanged when we excluded studies where SDs were estimated from other measures of uncertainty (e.g. standard error, 95% confidence interval) (Table 2).

The least precise MCIDs, which were based upon mean change SDs for the ADAS-Cog in patients randomized to receive memantine, were derived from only three RCTs and the pooled SD was influenced by one study (Table 3) [19]. When this latter study was removed, the pooled SD decreased to 6.8 and the MCIDs at 0.4 and 0.5 SDs were 2.7 and 3.4 points, respectively.

Discussion

We demonstrated how a distribution-based approach using systematic review methods can estimate MCIDs for scales reporting an outcome of interest. We found that our distribution-based approach derived MCIDs that were similar to accepted MCIDs for measuring changes in cognitive function in persons with Alzheimer disease [6, 20]. However, MCIDs derived from baseline scale score SDs were more precise than MCIDs derived from mean change scale score SDs, perhaps because mean change scale score SDs are dependent on baseline values and there are potential ceiling and floor effects associated with scales [21]. Furthermore, the least precise MCIDs were derived from a pooled estimate based on only three RCTs; therefore, deriving MCIDs from few studies may be less precise. We demonstrated how the pooled SD based upon only three RCTs was influenced by one study. When this study was removed, the MCIDs were similar to MCIDs in our primary analysis. The distribution-based method could be used where MCIDs for an outcome measure are not available; our approach could enhance knowledge user understanding of study results and facilitate planning of future studies through assistance with sample size calculation.

Our derived ADAS-Cog and MMSE MCIDs are similar to published MCIDs (Table 4) [6, 20, 22]. Using an anchor-based method, Schrag et al., found that persons with Alzheimer disease who had clinically important worsening on any of four anchor questions (memory, non-memory cognitive function, Functional Activities Questionnaire and Clinical Dementia Rating Scale) had a change in ADAS-Cog score of 2.7 to 3.8 points [22]. When Schrag et al., implemented a distribution-based method to estimate MCIDs at 0.5 SDs (using baseline ADAS-Cog score SDs), MCIDs ranged from 3.3 to 4.9 points for participants with a clinically meaningful decline on anchor questions [22]. Using an anchor-based approach, Rockwood et al., compared changes on the ADAS-Cog to clinician's interview based impression of change-plus caregiver input scores, patient/carer-goal attainment scaling, and clinician-goal attainment scaling. Rockwood et al., found that a change of 4 points on the ADAS-Cog was clinically important for persons with Alzheimer disease [20]. Our derived range of MCIDs for the ADAS-Cog encompasses these published MCIDs. Similarly, investigators from the DOMINO trial agreed that the MCID for a change in MMSE was 1 to 2 points among persons with Alzheimer disease [6]. Using a distribution-based approach, they estimated similar MCIDs for changes in MMSE scores, which ranged from 1.4 (assuming a distribution of 0.4 SDs) to 1.7 (assuming a distribution of 0.5 SDs) points [6]. Our derived range of MCIDs for the MMSE encompasses these published MCIDs as well. In contrast, using a survey of clinicians'

Table 1 Primary Analysis: MCIDs for Two Measures of Cognitive Function

	# RCTs (# Participants)	SD Range	Pooled SD	MCID: 0.4 x SD	MCID: 0.5 x SD
<i>MMSE</i>					
Baseline SDs	51 (12449)	0.94 to 6.80	4.0	1.6	2.0
Mean Change SDs	36 (10575)	0.33 to 6.12	3.6	1.4	1.8
<i>ADAS-Cog</i>					
Baseline SDs	37 (10006)	2.55 to 17.30	10.0	4.0	5.0
Mean Change SDs	38 (13288)	1.32 to 12.85	6.4	2.6	3.2

Abbreviations: Alzheimer Disease Assessment Scale – Cognitive Subscale (ADAS-Cog), minimum clinically important difference (MCID), Mini-Mental State Exam (MMSE), randomized trial (RCT), standard deviation (SD), number (#)

opinions, Burbach et al., found a MMSE MCID of 3.72 (95% confidence interval 3.50 to 3.95) points [12]. However, pooled SDs estimated from baseline and mean change MMSE scores in our meta-analysis were 4 and 3.6 (Table 1) points, respectively [14]; a MCID of 3.72 points represents a very large effect size [9, 14].

There are advantages to deriving MCIDs using systematic review methods and a distribution-based approach. Systematic reviews use explicit methods to synthesize evidence, which minimizes bias in the derivation of effect estimates and their associated measure of uncertainty [23]. Systematic reviews facilitate the generalization of results beyond any one study [23]. This is particularly important in the estimation of a MCID using our proposed distribution-based approach because a MCID is meant to be applied across a broad range of clinical scenarios. As demonstrated in our results, there is substantial variability in the distribution of uncertainty across individual studies. In general, systematic reviews also improve the accuracy of conclusions about the efficacy or safety of an intervention across study settings, which is why MCIDs derived with similar methods could also improve accuracy. Our proposed distribution-based approach could help knowledge users to assess whether an intervention has an effect on the outcome of interest over a range of clinically meaningful values (0.4 to 0.5 SDs), but researchers should be careful to select a

validated scale for measuring their outcome of interest [11, 24, 25].

If an outcome in a meta-analysis is reported with more than one scale, the pooled standard deviation (SD_{pooled}) estimated from systematic review data can also facilitate back-transformation of standardized mean differences derived from meta-analyses to mean differences. To derive a mean difference (MD_j) from a standardized mean difference (SMD_j), multiply SD_{pooled} by each standardized mean difference (SMD_j), as follows: $MD_j = SMD_j \times SD_{pooled}$. Researchers often either interpret a standardized mean difference with respect to thresholds first proposed by Cohen (i.e. 0.2 SDs represented a small difference and 0.8 SDs represented a large difference) or they back-transform standardized mean differences to mean differences, as described in the Cochrane Handbook of Systematic Reviews for Interventions [9, 17]. The Cochrane Handbook suggests using this method or SDs derived from an observational study related to the systematic review topic [17]. While observational data may be reflective of a real-world distribution of effect sizes, there are various biases that systematic reviewers must consider when deciding on which observational study to use, specifically, indication bias associated with comparing an intervention group to a non-intervention group in observational studies of interventions [26]. The influence of biases on a pooled SD (and their impact on

Table 2 Sensitivity Analysis: MCIDs for Two Measures of Cognitive Function

	# RCTs (# Participants)	SD Range	Pooled SD	MCID: 0.4 x SD	MCID: 0.5 x SD
<i>MMSE</i>					
Baseline SDs	38 (9614)	1.30 to 6.80	4.0	1.6	2.0
Mean Change SDs	12 (5288)	0.33 to 4.34	3.5	1.4	1.8
<i>ADAS-Cog</i>					
Baseline SDs	26 (5744)	4.60 to 17.3	9.9	4.0	5.0
Mean Change SDs	8 (3320)	1.32 to 7.88	6.4	2.6	3.2

Abbreviations: Alzheimer Disease Assessment Scale – Cognitive Subscale (ADAS-Cog), minimum clinically important difference (MCID), Mini-Mental State Exam (MMSE), randomized trial (RCT), standard deviation (SD), number (#)

Table 3 Secondary Analysis by Intervention Group: MCIDs for Two Measures of Cognitive Function

	# RCTs (# Participants)	SD Range	Pooled SD	MCID: 0.4 x SD	MCID: 0.5 x SD
Donepezil					
<i>MMSE</i>					
Baseline SDs	35 (3785)	1.08 to 5.90	4.2	1.7	2.1
Mean Change SDs	28 (3125)	0.33 to 6.12	3.6	1.4	1.8
<i>ADAS-Cog</i>					
Baseline SDs	22 (1693)	6.56 to 15.8	10.2	4.1	5.1
Mean Change SDs	20 (2215)	3.96 to 7.46	5.8	2.3	2.9
Galantamine					
<i>MMSE</i>					
Baseline SDs	7 (1285)	1.92 to 4.12	3.9	1.6	2.0
Mean Change SDs	5 (1102)	2.24 to 4.05	3.9	1.6	2.0
<i>ADAS-Cog</i>					
Baseline SDs	16 (2296)	5.02 to 11.78	9.7	3.9	4.9
Mean Change SDs	22 (3179)	5.00 to 7.43	6.0	2.4	3.0
Rivastigmine					
<i>MMSE</i>					
Baseline SDs	17 (1944)	0.98 to 4.9	3.5	1.4	1.8
Mean Change SDs	12 (1891)	0.46 to 3.6	3.2	1.3	1.6
<i>ADAS-Cog</i>					
Baseline SDs	14 (1825)	4.60 to 12.30	10.0	4.0	5.0
Mean Change SDs	15 (2892)	1.32 to 12.85	7.1	2.8	3.6
Memantine					
<i>MMSE</i>					
Baseline SDs	9 (548)	1.60 to 6.20	4.0	1.6	2.0
Mean Change SDs	4 (442)	2.20 to 5.65	4.1	1.6	2.1
<i>ADAS-Cog</i>					
Baseline SDs	5 (706)	7.90 to 11.01	10.0	4.0	5.0
Mean Change SDs	3 (603)	5.46 to 9.77	8.2	3.3	4.1
Placebo					
<i>MMSE</i>					
Baseline SDs	36 (4396)	0.94 to 6.8	4.1	1.6	2.1
Mean Change SDs	27 (3758)	0.33 to 5.76	3.7	1.5	1.9
<i>ADAS-Cog</i>					
Baseline SDs	28 (3398)	2.55 to 17.3	10.1	4.1	5.1
Mean Change SDs	29 (4315)	2.50 to 8.19	6.3	2.5	3.2

Abbreviations: Alzheimer Disease Assessment Scale – Cognitive Subscale (ADAS-Cog), minimum clinically important difference (MCID), Mini-Mental State Exam (MMSE), randomized trial (RCT), standard deviation (SD), number (#)

derived mean differences) derived from RCTs included in a systematic review can be tested in sensitivity analyses, which can increase confidence in findings.

There are limitations to using our proposed distribution-based approach. It is unclear if MCIDs generated by this approach are generalizable to all situations in which a scale is used. For example, MCIDs derived

from the systematic review and network meta-analysis of the comparative effectiveness and safety of cognitive enhancers (cholinesterase inhibitors and memantine) for treating Alzheimer disease might not be generalizable to MCIDs for these scales if using a nonpharmacologic intervention (e.g. exercise, cognitive training); however, MCIDs for determining meaningful changes in pain

Table 4 Comparison of Derived to Published MCIDs

	Published MCIDs	Derived MCID: 0.4 x SD	Derived MCID: 0.5 x SD
<i>MMSE</i>			
Baseline SDs	1.7 to 2.1 [6]	1.6	2.0
Mean Change SDs	1 to 2 [6]; 1.4 to 1.7 [6]; 3.72 [12]	1.4	1.8
<i>ADAS-Cog</i>			
Baseline SDs	3.3 to 4.9 [22]	4.0	5.0
Mean Change SDs	2.7 to 3.8 [22]; 4 [20]	2.6	3.2

Abbreviations: Alzheimer Disease Assessment Scale – Cognitive Subscale (ADAS-Cog), minimum clinically important difference (MCID), Mini-Mental State Exam (MMSE), standard deviation (SD)

scores for patients with osteoarthritis did not vary across pharmacologic (nonsteroidal anti-inflammatories), non-pharmacologic (i.e. rehabilitation), or surgical (i.e. total hip replacement, total knee replacement) interventions [27]. And, similar to other distribution-based approaches, the anticipated distribution of uncertainty may vary based on effect modifiers; therefore, it will be important to consider a plausible distribution of values for the MCID (i.e. 0.4 to 0.5 SDs) when interpreting results [6, 9, 10]. These limitations will need to be explored in future studies.

Conclusion

We demonstrated how a distribution-based approach using systematic review data can estimate MCIDs for scale-based outcomes in a systematic review of interventions. Given that MCIDs represent thresholds for clinically discernible changes as measured on a scale, it is important for researchers to have a way of estimating MCIDs for outcomes derived from systematic reviews that can be communicated with knowledge users. We believe this distribution-based approach will help knowledge users to better understand the clinical importance of outcomes reported in systematic reviews and meta-analyses and it can estimate MCIDs where no published estimates exist, thereby facilitating shared decision making. Future research should focus on the generalizability of this method to other clinical settings by using scale-based outcome measures from systematic reviews of RCTs of interventions in other healthcare disciplines. Our method could also be used in the design of future trials of interventions to estimate sample sizes required to show clinically meaningful differences for patients and to help patients and clinicians interpret trial outcomes.

Abbreviations

ADAS-Cog: Alzheimer Disease Assessment Scale – Cognitive Subscale; MD: Mean difference; MMSE: Mini-Mental State Exam; MCID: Minimum clinically important difference; NMA: Network meta-analysis; SD_{pooled}: Pooled standard deviation; RCT: Randomized trial; SD: Standard deviation; SMD: Standardized mean difference

Acknowledgements

We would like to thank the following people who contributed to data curation in the original publication from which we retrieved our datasets: Huda M. Ashoor, Charlene Soobiah, Patricia Rios, John D. Ivory, Paul A. Khan, Fatemeh Yazdi, Marco Ghassemi, Erik Blondal, Joanne M. Ho, and Carmen H. Ng [14].

Data statement

Available on reasonable request from Dr. Straus (e-mail, sharon.straus@utoronto.ca).

Authors' contributions

JAW, AAV and SES contributed to the design of this study. ACT and SES provided access to study data. JAW and AAV conducted data analyses. JAW drafted the first version of this manuscript and all authors contributed to its revision. The authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. ACT is funded by a Tier 2 Canada Research Chair in Knowledge Synthesis. SES is funded by a Tier 1 Canada Research Chair in Knowledge Translation and the Squires Chalmer Chair in Medicine. AAV is funded by a European Union's Horizon 2020 grant [No. 754936].

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Authors have no competing interests to declare.

Author details

¹Knowledge Translation Program, Li Ka Shing Knowledge Institute, St. Michael's Hospital- Unity Health Toronto, 209 Victoria Street, East Building, Room 723, Toronto, Ontario M5B 1W8, Canada. ²Division of Geriatric Medicine, Department of Medicine, University of Toronto, 190 Elizabeth Street, R. Fraser Elliott Building, 3-805, Toronto, Ontario M5G 2C4, Canada. ³Department of Primary Education, School of Education, University of Ioannina, 45110 Ioannina, Greece. ⁴Department of Surgery & Cancer, Faculty of Medicine, Imperial College, Institute of Reproductive and Developmental Biology, W12 0NN, London, UK. ⁵Institute for Health Policy, Management and Evaluation, University of Toronto, 4th floor, 155 College St, Toronto, Ontario M5T 3M6, Canada.

Received: 23 November 2020 Accepted: 2 February 2021

Published online: 26 February 2021

References

1. Lemstra AW, Richard E, van Gool WA. Cholinesterase inhibitors in dementia: yes, no, or maybe? *Age Ageing*. 2007;36(6):625–7.

2. Pelosi AJ, McNulty SV, Jackson GA. Role of cholinesterase inhibitors in dementia care needs rethinking. *BMJ*. 2006;333(7566):491–3.
3. McGlothlin AE, Lewis RJ. Minimal clinically important difference. *JAMA*. 2014;312(13):1342–3.
4. Copay AG, Subach BR, Glassman SD, Polly DW Jr, Schuler TC. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J*. 2007;7(5):541–6.
5. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. *Control Clin Trials*. 1989;10(4):407–15.
6. Howard R, Phillips P, Johnson T, O'Brien J, Sheehan B, Lindesay J, Bentham P, Burns A, Ballard C, Holmes C, et al. Determining the minimum clinically important differences for outcomes in the DOMINO trial. *Int J Geriatr Psychiatry*. 2011;26(8):812–7.
7. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J Clin Epidemiol*. 2017;82:128–36.
8. Henderson EJ, Morgan GS, Amin J, Gaunt DM, Ben-Shlomo Y. The minimum clinically important difference (MCID) for a falls intervention in Parkinson's: a delphi study. *Parkinsonism Relat Disord*. 2019;61:106–10.
9. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
10. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life. *Med Care*. 2003;41:582–92.
11. Ballard C, Banister C, Khan Z, Cummings J, Demos G, Coate B, Youakim JM, Owen R, Stankovic S, Tomkinson EB, et al. Evaluation of the safety, tolerability, and efficacy of pimavanserin versus placebo in patients with Alzheimer's disease psychosis: a phase 2, randomised, placebo-controlled, double-blind study. *The Lancet Neurology*. 2018;17(3):213–22.
12. Burback D, Molnar FJ, St. John P, Man-Son-Hing M. Key methodological features of randomized trials of Alzheimer's disease therapy. *Dement Geriatr Cogn Disord*. 1999;10:534–40.
13. Redelmeier DA, Kahneman D. Patients' memories of painful medical treatments: real-time and retrospective evaluations of two minimally invasive procedures. *Pain*. 1996;66:3–8.
14. Tricco AC, Ashoor HM, Soobiah C, Rios P, Veroniki AA, Hamid JS, Ivory JD, Khan PA, Yazdi F, Ghassemi M, et al. Comparative effectiveness and safety of cognitive enhancers for treating Alzheimer's disease: systematic review and network Metaanalysis. *J Am Geriatr Soc*. 2018;66(1):170–8.
15. Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State" A Practical Method for Grading the Cognitive State of Patients for the Clinician. *J psychiatr Res*. 1975;12:189–98.
16. Rosen WG, Mohs RC, Davis KL. A new rating scale for Alzheimer's disease. *Am J Psychiatr*. 1984;141(11):1356–64.
17. *Cochrane Handbook for Systematic Reviews of Interventions*. In: Edited by Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA: The Cochrane Collaboration; 2020.
18. Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *J Clin Epidemiol*. 2006;59(1):7–10.
19. Bakchine S, Loft H. Memantine treatment in patients with mild to moderate Alzheimer's disease: results of a randomised, double-blind, placebo-controlled 6-month study. *J Alzheimers Dis*. 2008;13:97–107.
20. Rockwood K, Fay S, Gorman M. The ADAS-cog and clinically meaningful change in the VISTA clinical trial of galantamine for Alzheimer's disease. *Int J Geriatr Psychiatry*. 2010;25(2):191–201.
21. Hays RD, Woolley JM. The concept of clinically meaningful difference in health-related quality-of-life research: how meaningful is it? *Pharmacoeconomics*. 2000;18:419–23.
22. Schrag A, Schott JM. Alzheimer's disease neuroimaging I: what is the clinically relevant change on the ADAS-cog? *J Neurol Neurosurg Psychiatry*. 2012;83(2):171–3.
23. Mulrow CD. Rationale for systematic reviews. *BMJ*. 1994;309:597–9.
24. Mokkink LB, Boers M, van der Vleuten CPM, Bouter LM, Alonso J, Patrick DL, de Vet HCW, Terwee CB. COSMIN risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: a Delphi study. *BMC Med Res Methodol*. 2020;20(1):293.
25. Santesso N, Barbara AM, Kamran R, Akkinepally S, Cairney J, Akl EA, Schunemann HJ. Conclusions from surveys may not consider important biases: a systematic survey of surveys. *J Clin Epidemiol*. 2020;122:108–14.
26. Jepsen P, Johnsen SP, Gillman MW, Sorensen HT. Interpretation of observational studies. *Heart*. 2004;90(8):956–60.
27. Katz NP, Paillard FC, Ekman E. Determining the clinical importance of treatment benefits for interventions for painful orthopedic conditions. *J Orthop Surg Res*. 2015;10:24.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

