EDITORIAL

Seeing the patient's perspective: a guide to patient-reported outcome measures and minimal important differences in ophthalmic research

© The Author(s), under exclusive licence to The Royal College of Ophthalmologists 2022

Eye (2022) 36:1339–1341; <https://doi.org/10.1038/s41433-022-01989-5>

INTRODUCTION

In recent years, emphasis on patient-centered care has led to increased use of patient-reported outcomes (PROs) in health research. A PRO is “any report of the status of a patient’s health condition that comes directly from the patient, without interpretation of the patient’s response by a clinician or anyone else” [1]. PROMs are instruments that measure PROs [2]. Trialists have used vision-related PROMs, such as the National Eye Institute-Visual Function Questionnaire-25 (NEI-VFQ-25), as either primary or secondary endpoints [3–5]. Although PROMs provide important information about the efficacy of interventions and the impact of disease on patient health, interpretation of PROM data has inherent challenges. In this editorial, we discuss the NEI-VFQ-25, an example of a commonly used PROM in ophthalmology trials, and the concept of minimal important difference (MID) that is frequently used to enhance the interpretability of PROM data.

THE NATIONAL EYE INSTITUTE-VISUAL FUNCTION QUESTIONNAIRE (NEI-VFQ)

The NEI-VFQ-25 measures the impact of visual impairment on health-related quality of life [6–8]. Each item of the PROM measures impairment, activity limitation, or participant restriction, following the World Health Organization’s International Classification of Functioning Disability and Health framework for measuring the health-related impact of disease [6]. Mangione et al. developed the original 51-item NEI-VFQ in a cohort of patients with age-related cataracts, age-related macular degeneration, diabetic retinopathy, primary open-angle glaucoma, cytomegalovirus retinitis, or low vision from any cause [7, 8]. Subsequently, the group shortened the NEI-VFQ to 25-item, 9-item, and 8-item versions of the PROM [9–11].

The NEI-VFQ-25 has three domains—general health and vision (4 items), difficulty with activities (12 items), and responses to vision problems (9 items)—that include items related to near vision, distance vision, driving, peripheral vision, ocular pain, role limitations, dependency, social function, and mental health. Scores for the NEI-VFQ-25 range from 0 to 100, where 100 represents the best possible score. The NEI-VFQ-25, a reliable and validated instrument, can be interviewer- or self-administered, and is translated into over 30 languages [12].

THE CHALLENGE OF INTERPRETING PRO DATA

Consider a clinical trial that compares a new anti-VEGF agent for neovascular age-related macular degeneration with a currently approved anti-VEGF agent. The objective of this trial is to demonstrate non-inferiority of the new agent to the current standard of care in terms of visual acuity in the study eye. Trialists administer the NEI-VFQ-25 at baseline and at the primary endpoint. What if the trial reports non-inferiority in visual acuity at one year and a mean difference of 6.3 points ($p < 0.05$) in NEI-VFQ-25 scores between trial arms? How should clinicians interpret these results?

Although the mean difference in PROM score is statistically significant, it is unclear whether a difference of 6.3 points on the NEI-VFQ-25 is trivial, or if this difference is meaningful to patients. Given this challenge, an MID can be utilized to facilitate the interpretation of NEI-VFQ-25 scores.

WHAT IS AN MID AND HOW DO WE ASSESS ITS CREDIBILITY?

In 1989, Jaeschke et al. coined the term MID, defined as the smallest difference in outcome that patients perceive as important [13]. Two approaches are widely used to estimate MIDs: distribution-based and anchor-based methods [14]. To calculate an MID using a distribution-based method (e.g., 0.5 standard deviation method [15]), researchers rely on statistical characteristics of the study sample, the measurement properties of the PROM, or the statistical significance of the observed change in PROM score. As such, MIDs estimated using distribution-based methods do not necessarily represent differences in PROM scores that are meaningful to patients. Rather, distribution-based MIDs identify a difference in PROM score that is unlikely to be due to random measurement error [16]. In contrast, when using the anchor-based approach to calculate MIDs, researchers relate changes in PROM score to another meaningful external criterion (or anchor) such as a transition item (e.g., How is your vision now compared with how it was before your cataract surgery: much worse, a little worse, same, a little better, or much better? [17]). Thus, when using anchor-based methods to estimate an MID, researchers ascribe meaning to observed changes in PROM score.

Although anchor-based MIDs are preferred to distribution-based MIDs, some anchor-based MID estimates may be more (or less) trustworthy than other MID estimates. The trustworthiness of an anchor-based MID is limited by the choice of anchor—its validity, reliability, and relationship with the PROM—and the statistical method used to calculate the MID [16, 18].

There are five key criteria that researchers may use to assess the credibility of an anchor-based MID [18]:

1. Did the patient respond directly to the PROM and the anchor?
2. Is the anchor easily understandable and relevant for patients?
3. Has the anchor shown a good correlation with the PROM?
4. Is the MID precise?
5. Does the threshold or difference between groups on the anchor used to estimate the MID reflect a small but important difference?

It is imperative that researchers, clinicians, and decision makers using PROMs be able to identify MIDs with high credibility; use of an untrustworthy MID for sample size estimation may undermine the results of an otherwise well-designed RCT or lead to the flawed interpretation of PROM data in trials and systematic reviews.

MIDS FOR THE NEI-VFQ-25

We queried PROMID (www.promid.org), a living inventory of anchor-based MIDs, for PROMs evaluated in patients with ophthalmic conditions [19]. As of September 2021, researchers reported 128 anchor-based MID estimates, including MIDs for the Visual Function Index (28%; 36/128) and the NEI-VFQ-25 (45%; 58/128). Fifty-eight MIDs estimated for the NEI-VFQ-25 are reported in four primary publications: three studies in patients with glaucoma [3, 20, 21], and the fourth study in patients with intermediate or posterior uveitis [22].

Overall, available MIDs for the NEI-VFQ-25 are of low credibility. These estimates are greatly limited by the researchers' choice of anchor and threshold used to estimate an MID, failure to report the correlation between the PROM and the anchor, and imprecision of MID estimates. Except for Burr et al. who used the EQ-5D-3L as an anchor [21], researchers typically estimated MIDs using the best-corrected visual acuity score [3, 20, 22] or the mean deviation from the visual field [3, 20] as an anchor. Thus, patients did not directly respond to the anchor, nor was the anchor judged to be intuitively understandable for patients. In two studies (50%; 29/58 MIDs) researchers failed to report the correlation between the PROM and the anchor, while the third group of researchers reported a correlation of ≥ 0.3 (43%; 25/58 MIDs). Such insufficient reporting limits assessors' ability to make judgments about the relationship between the PROM and anchor, which is critical for establishing credibility. In most cases, authors reported measures of precision associated with the MID, albeit with marked variability in point estimates. Finally, despite the goal to determine an MID, the thresholds on the anchor used to estimate the MID did not actually reflect a difference that is small in magnitude yet important to patients.

To calculate high credibility MIDs for the NEI-VFQ-25, we recommend that researchers use an anchor that is easily understandable and relevant to patients, and ensure the PROM and anchor measure closely related concepts. To demonstrate the validity of the anchor, the correlation between the PROM and anchor should be at least 0.5 [23]. Researchers should also select a threshold on the anchor that reflects an MID, and ideally, provide empirical evidence to justify their choice of threshold.

CONCLUSION

Researchers considering the use of PROs in evidence-based decision making should ensure that PROMs have adequate measurement properties and are appropriate for the clinical context in which they are used. MIDs are valuable in enhancing the interpretation of PRO-based evidence from individual trials and meta-analyses, and sample size estimation. Despite the widespread use and reporting of vision-related PROMs such as the

NEI-VFQ-25, interpretation of these data should be performed with caution as high credibility MIDs are lacking. Researchers estimating MIDs for the NEI-VFQ-25 should select easily interpretable, patient-important anchors with a strong correlation with the PROM, and ensure that the threshold selected to estimate an MID is appropriate.

Copyright © 2018 McMaster University, Hamilton, Ontario, Canada.

The Minimal Important Difference Credibility Assessment Tool and Minimal Important Difference Inventory, authored by Dr Tahira Devji et al., is the copyright of McMaster University (Copyright © 2018, McMaster University, Hamilton, Ontario, Canada). The Minimal Important Difference Credibility Assessment Tool and Minimal Important Difference Inventory have been provided under license from McMaster University and must not be copied, distributed or used in any way without the prior written consent of McMaster University.

Contact the McMaster Industry Liaison Office at McMaster University, email: milo@mcmaster.ca for licensing details.

Anila Qasim¹, Tahira Devji^{1,2}, Mark R. Phillips¹, Charles C. Wykoff^{3,4}, Peter K. Kaiser⁵, Lehana Thabane^{1,6}, Mohit Bhandari^{1,7}, Varun Chaudhary^{1,7} and for the Retina Evidence Trials InterNational Alliance (R.E.T.I.N.A.) Study Group*
¹Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada. ²Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada. ³Retina Consultants of Texas (Retina Consultants of America), Houston, TX, USA. ⁴Blanton Eye Institute, Houston Methodist Hospital, Houston, TX, USA. ⁵Cole Eye Institute, Cleveland Clinic, Cleveland, OH, USA. ⁶Biostatistics Unit, St. Joseph's Healthcare-Hamilton, Hamilton, ON, Canada. ⁷Department of Surgery, McMaster University, Hamilton, ON, Canada. *A list of authors and their affiliations appears at the end of the paper. [✉]email: vchaudh@mcmaster.ca

REFERENCES

1. Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. U.S. Department of Health and Human Services FDA; 2009.
2. Johnston B, Patrick D, Devji T, Maxwell L, III CB, Beaton D, et al. Chapter 18: patient-reported outcomes. *Cochrane Handbook for Systematic Reviews of Interventions*. Cochrane. 2021. www.training.cochrane.org/handbook.
3. Kotecha A, Feuer WJ, Barton K, Gedde SJ. Quality of life in the tube versus trabeculectomy study. *Am J Ophthalmol*. 2017;176:228–35.
4. Mitchell P, Massin P, Bressler S, Coon CD, Petrillo J, Ferreira A, et al. Three-year patient-reported visual function outcomes in diabetic macular edema managed with ranibizumab: the RESTORE extension study. *Curr Med Res Opin*. 2015;31:1967–75.
5. Bressler NM, Chang TS, Suñer IJ, Fine JT, Dolan CM, Ward J, et al. Vision-related function after ranibizumab treatment by better- or worse-seeing eye: clinical trial results from MARINA and ANCHOR. *Ophthalmology*. 2010;117:747–56.e4.
6. Nassiri N, Mehravaran S, Nouri-Mahdavi K, Coleman AL. National Eye Institute Visual Function Questionnaire: usefulness in glaucoma. *Optom Vis Sci*. 2013;90:745–53.
7. Mangione CM, Lee PP, Pitts J, Gutierrez P, Berry S, Hays RD. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). NEI-VFQ Field Test Investigators. *Arch Ophthalmol*. 1998;116:1496–504.
8. Parrish RK II. Visual impairment, visual functioning, and quality of life assessments in patients with glaucoma. *Trans Am Ophthalmol Soc*. 1996;94:919–1028.
9. Coleman AL, Yu F, Ensrud KE, Stone KL, Cauley JA, Pedula KL, et al. Impact of age-related macular degeneration on vision-specific quality of life: follow-up from the 10-year and 15-year visits of the Study of Osteoporotic Fractures. *Am J Ophthalmol*. 2010;150:683–91.
10. Kodjebacheva G, Coleman AL, Ensrud KE, Cauley JA, Yu F, Stone KL, et al. Reliability and validity of abbreviated surveys derived from the National Eye Institute Visual Function Questionnaire: the study of osteoporotic fractures. *Am J Ophthalmol*. 2010;149:330–40.
11. Mangione CM, Lee PP, Gutierrez PR, Spritzer K, Berry S, Hays RD. Development of the 25-item National Eye Institute Visual Function Questionnaire. *Arch Ophthalmol*. 2001;119:1050–8.

12. Trust MR. National Eye Institute Visual Function Questionnaire-25 (NEI-VFQ-25). 2021. <https://eprovide.mapi-trust.org/instruments/national-eye-institute-visual-function-questionnaire-25>.
13. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10:407–15.
14. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56:395–407.
15. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582–92.
16. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. *JAMA*. 2014;312:1342–3.
17. Quintana JM, Aguirre U, Las-Hayas C, Gonzalez N, Garcia S, Escobar A. Use of the patient acceptable symptom state and the minimal clinically important difference to evaluate the outcomes of cataract extraction. *Am J Ophthalmol*. 2011;152:234–43.e3.
18. Devji T, Carrasco-Labra A, Qasim A, Phillips M, Johnston BC, Devasenapathy N, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. *BMJ*. 2020;369:m1714.
19. Carrasco-Labra A, Devji T, Qasim A, Phillips MR, Wang Y, Johnston BC, et al. Minimal important difference estimates for patient-reported outcomes: a systematic survey. *J Clin Epidemiol*. 2021;133:61–71.
20. Gillespie BW, Musch DC, Niziol LM, Janz NK. Estimating minimally important differences for two vision-specific quality of life measures. *Invest Ophthalmol Vis Sci*. 2014;55:4206–12.
21. Burr JM, Cooper D, Ramsay CR, Che Hamzah J, Azuara-Blanco A. Interpretation of change scores for the National Eye Institute Visual Function Questionnaire-25: the minimally important difference. *Br J Ophthalmol*. 2021. E-pub ahead of print.
22. Naik RK, Gries KS, Rentz AM, Kowalski JW, Revicki DA. Psychometric evaluation of the National Eye Institute Visual Function Questionnaire and Visual Function Questionnaire Utility Index in patients with non-infectious intermediate and posterior uveitis. *Qual Life Res*. 2013;22:2801–8.
23. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. *J Clin Epidemiol*. 2002;55:900–8.

ACKNOWLEDGEMENTS

We thank Elena Kum for proofreading an early draft of this editorial.

FOR THE RETINA EVIDENCE TRIALS INTERNATIONAL ALLIANCE (R.E.T.I.N.A.) STUDY GROUP

Varun Chaudhary^{1,7}✉, Mohit Bhandari^{1,7}, Charles C. Wykoff^{2,3,4}, Sobha Sivaprasad⁸, Lehana Thabane^{1,6}, Peter Kaiser⁵, David Sarraf⁹, Sophie J. Bakri¹⁰, Sunir J. Garg¹¹, Rishi P. Singh^{12,13}, Frank G. Holz¹⁴, Tien Y. Wong^{15,16} and Robyn H. Guymer^{17,18}

⁸NIHR Moorfields Biomedical Research Centre, Moorfields Eye Hospital, London, UK. ⁹Retinal Disorders and Ophthalmic Genetics, Stein Eye Institute, University of California, Los Angeles, CA, USA. ¹⁰Department of Ophthalmology, Mayo Clinic, Rochester, MN, USA. ¹¹The Retina Service at Wills Eye Hospital, Philadelphia, PA, USA. ¹²Center for Ophthalmic Bioinformatics, Cole Eye Institute, Cleveland Clinic, Cleveland, OH, USA. ¹³Cleveland Clinic Lerner College of Medicine, Cleveland, OH, USA. ¹⁴Department of Ophthalmology, University of Bonn, Bonn, Germany. ¹⁵Singapore Eye Research Institute, Singapore, Singapore. ¹⁶Singapore National Eye Centre, Duke-NUS Medical School, Singapore, Singapore. ¹⁷Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, VIC, Australia. ¹⁸Department of Surgery (Ophthalmology), The University of Melbourne, Melbourne, VIC, Australia.

AUTHOR CONTRIBUTIONS

AQ was responsible for writing, critical review, and feedback on manuscript. TD was responsible for writing, critical review, and feedback on manuscript. MRP was responsible for conception of idea, critical review, and feedback on manuscript. CCW was responsible for critical review and feedback on manuscript. PKK was responsible for critical review and feedback on manuscript. LT was responsible for critical review and feedback on manuscript. MB was responsible for conception of idea, critical review, and feedback on manuscript. VC was responsible for conception of idea, critical review, and feedback on manuscript.

COMPETING INTERESTS

AQ, TD, and MRP: nothing to disclose. CCW: consultant: Acuela, Adverum Biotechnologies, Inc., Aerpio, Alimera Sciences, Allegro Ophthalmics, LLC, Allergan, Apellis Pharmaceuticals, Bayer AG, Chengdu Kanghong Pharmaceuticals Group Co, Ltd, Clearside Biomedical, DORC (Dutch Ophthalmic Research Center), EyePoint Pharmaceuticals, Gentech/Roche, GyroscopeTx, IVERIC bio, Kodiak Sciences Inc., Novartis AG, ONL Therapeutics, Oxurion NV, PolyPhotonix, Recens Medical, Regeneron Pharmaceuticals, Inc., REGENXBIO Inc., Santen Pharmaceutical Co, Ltd, and Takeda Pharmaceutical Company Limited; research funds: Adverum Biotechnologies, Inc., Aerpio Pharmaceuticals, Inc., Aerpio, Alimera Sciences, Allergan, Apellis Pharmaceuticals, Chengdu Kanghong Pharmaceutical Group Co, Ltd, Clearside Biomedical, Gemini Therapeutics, Genentech/Roche, Graybug Vision, Inc., GyroscopeTx, Ionis Pharmaceuticals, IVERIC bio, Kodiak Sciences Inc., Neurotech LLC, Novartis AG, Opthea, Outlook Therapeutics, Inc., Recens Medical, Regeneron Pharmaceuticals, Inc., REGENXBIO Inc., Samsung Pharm Co, Ltd, Santen Pharmaceutical Co, Ltd, and Xbrane Biopharma AB—unrelated to this study. PKK: consultant: Novartis, Bayer, Biogen, Regeneron, Kanghong, Allergan, RegenxBio—unrelated to this study. LT: nothing to disclose. MB: research funds: Pendopharm, Bioventus, Acumed—unrelated to this study. VC: advisory board member: Alcon, Roche, Bayer, Novartis; Grants: Bayer, Novartis—unrelated to this study.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Varun Chaudhary.

Reprints and permission information is available at <http://www.nature.com/reprints>

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