GUEST EDITORIAL

Vision-Related Quality of Life

Since the World Health Organization (WHO) stated that "health is a state of complete physical, mental, and social well-being and not merely the absence of disease or infirmity," the impact of a disease or treatment on the patient's life has become increasingly measured using assessments of well-being or quality of life (QoL) in addition to more traditional clinical measures. Indeed, health-related QoL (HR-QoL) measurements are now a standard part of clinical trials of health interventions. Although there is some debate about the precise definition of HR-QoL,[1] the content of HR-QoL instruments typically includes an assessment of the ability to perform activities of daily living, interactions with other people, emotional well-being, and independence. Other dimensions have also been measured.[2]

Because of the breadth of HR-QoL and its patient-centered nature, it has been measured using questionnaires (called instruments in the research literature), which can be efficient tools for gathering large amounts of data quickly. A prime example of the importance of these instruments in measuring outcome is the evaluation of the usefulness of second eye cataract surgery. In the 1990s, several U.S. insurance companies and UK health authorities suggested that cataract surgery for the second eye was not necessary based on the relatively minor benefits provided in terms of binocular visual acuity.[3] However, several studies used self-reported questionnaire data to provide evidence that patients experienced substantial improvements in performing everyday activities that are dependent on vision as well as in symptoms and QoL.[3–5] The value of the surgery is no longer questioned.

Most recently, there are concerns that many older adults with low vision that have cataract as a secondary diagnosis,[6] are not considered for surgery, in large part because studies have demonstrated only limited improvements in VA. Yet few studies have examined the impact of such surgery on patient QoL, independence, and community participation. Included in this special issue is a report that highlights the major improvements in VR-QoL that cataract surgery can make to patients with both cataract and age-related macular degeneration.[7] These examples highlight the value of QoL instruments in demonstrating the range of benefits possible from cataract surgery.

Symptoms and Quality of Life

The content of this feature issue is not confined to instruments that purport to measure global QoL. A number of instruments are described in this issue that measure symptoms including ocular pain,[8] visual and physical symptoms related to VDU use,[9] and visual symptoms in college students.[10] These studies reflect the significance of symptoms and their management within the optometric setting. This area of optometric research has previously lacked an evidence base due to the difficulties in quantifying symptoms. The papers in this Feature

Issue represent an exciting contribution that should facilitate more research in this core optometry area.

Quality of Life and Health Economics

QoL measures are also being used in health cost-effectiveness research to try to determine how economic resources can be best used in healthcare, and this area of research is reviewed by Kymes.[11] The currency often used in these studies are "quality-adjusted life years" (QALYs), which equal 1 for each year of full-health life gained, and <1 for various degrees of illness or disability. Thus the cost-effectiveness of a treatment can be assessed by the cost per QALY produced. To take the example used earlier, research has suggested that second eye cataract surgery costs $2,727 per QALY gained and is an extremely cost-effective procedure when compared with other interventions across medical specialties and only slightly less than first-eye cataract surgery ($2,023 /QALY gained).[12] In this feature issue, the development of a utility scale, which could be used to "quality adjust" life years, for use in glaucoma is reported.[13]

Likert (summary) Scales

VR-QoL questionnaires typically score responses using Likert scales. Likert scales are classically five point scales on which respondents signify their agreement to a statement.[14] For example, the statement "I feel embarrassed wearing my spectacles" could be responded to on a scale that includes "strongly agree", "somewhat agree", "neutral", "somewhat disagree" and "strongly disagree." A scale similar to this has been used to compare the QoL in young spectacle and contact lens wearers in the Psychosocial Impact of Assistive Devices Scale (PIADS) questionnaire.[15] As the use of instruments extended beyond psychology to medical fields, the format and purpose of questionnaires changed. Unfortunately, the change in the design (particularly the format of the response categories deviating from agreement with a statement) and application of the questionnaires also meant that traditional methods of simply adding up the values assigned to categories to form a total score became invalid, but this was not recognized for many of the early VR-QoL instruments. For example, the Activities of Daily Vision Scale (ADVS) questionnaire assessed activity limitation, such as driving at night on a Likert-type scale with responses of *no difficulty*, *a little difficulty*, *moderate difficulty*, *extreme difficulty* and *unable to do the activity because of their vision*.[16] Responses to a series of items are then summed to provide an overall ADVS score and scores for subscales such as driving, near vision, distance vision, etc. Summing scores in this way can only provide true measurement if all items (or questions) are of equal difficulty and if the magnitude

of the differences between the response categories are the same.[17] Unfortunately, summed scores from Likert-type scales are fairly prevalent among VR-QoL instruments at present and they do not provide true measurements as explained by Mallinson[18] in this issue. For example, driving at night and driving during the day, as used in the ADVS, clearly have different degrees of difficulty[17] and scores from each item should not have an equal weighting in a "visual activity limitation" score.[19]

Rasch Analysis

Rasch analysis[20] is a special case of Item Response Theory (IRT), whereby items and persons can be scaled according to a series of responses to items (See Massof's background paper[21] in this issue for further discussion of these methods when used to measure vision disability). Rasch analysis places items and persons on the same linear scale, ordering subjects from most able to the least able and items from most difficult to least difficult. Indeed, Rasch analysis and related IRT models have now penetrated the field, being applied in a majority of papers in this topical issue.[7,10,22,23] A number of papers have used Rasch analysis for testing instrument validity or applicability to a different population.[24–26] It is also critical when developing a questionnaire,[2] as highlighted by Court and colleagues[27] in this issue, as it can be used to identify the most useful items and also those that could be discarded on the basis that they either provide redundant or irrelevant information.

Future Trends

Rasch analysis can add a great deal to our understanding of how important constructs are operationalized in the hierarchical ordering of items and in how patients respond to items, and has a role in further understanding widely used questionnaires in the vision literature. Papers in this Feature Issue address this. Specifically, the publication of basic conversions to Rasch scoring will simplify clinical application of this important scoring method.[24,28] Newly developed questionnaires should utilize Rasch analysis or other IRT models in both their development and scoring.[18,19]

One of the unique features of IRT approaches is that they do not require that all respondents answer all questions in order to estimate either person ability or item difficulty. The utility of this feature is that respondents need only answer the questions that are best targeted to their level of function, improving measurement precision and reducing respondent burden. This feature has also enabled the development of computer-adaptive testing (CAT).[29] CAT is based on the development of items banks, which are large groups of items that cover a fairly broad range of the trait being measured. Items within these banks may be newly developed specifically for the bank, drawn from existing instruments that all measure the same construct, or a combination of these approaches.[30] The items are administered to a wide range of patients and the item bank is calibrated using item response theory approaches. It is this calibrated item bank that allows the implementation of CAT. CAT is a particular kind of computer administration in which each subsequent item is presented to a respondent based on his or her prior response.

IRT is also used in the development of "short form" instruments. Short forms are a preselected, fixed subset of items that represent the range of items in the bank and the same set of items are administered to all respondents. Short forms, selected from an item bank generally result in more precise person measures than existing instruments because IRT enables the selection of highly effective items.[31] CAT is very efficient since only a very small number of items need be presented to a respondent to get a very precise person measure. However, it requires access to sophisticated computer software and interfaces. Short forms often produce person measures that are almost as precise as CAT administrations but can be administered in pencil and paper format.

There are currently several major attempts to develop item banks. The most advanced is PROMIS – Patient Reported Outcome Measures Information System, funded by the National Institutes of Health. This multi-center study is developing item banks in the areas of pain, fatigue, emotional distress, physical functioning, and social role participation.[32] Efforts such as PROMIS are intended for use with multiple diagnostic groups. Other efforts include the development of a QoL item banks for specific conditions, such as persons with arthritis.[33] Other notable self-report item banking efforts include the Activity Measure for PostAcute Care (AM-PAC), which measures mobility, personal care, and functional cognition and can be administered in either CAT or short-form versions.[34]

While CAT has the advantages of comprehensive coverage of a trait, efficiency, and precision, one of the major advantages is the ability to interpret results across multiple studies. With so many QoL and functional performance scales it can be difficult to interpret findings across studies. If these multiple instruments are calibrated within the same item bank, in the same frame of reference, then results can be reported in the frame of reference, enabling easier comparison across studies. This Feature Issue provides insight into this future with an adaptive testing instrument.[35] Item banking of the many visual disability questionnaires remains a worthy goal of future research.

We are delighted to bring you this Feature Issue on *Vision-Related Quality of Life*, which includes a range of important research papers that highlight the usefulness of these outcome measures in ophthalmic research. Hopefully, this issue emphasizes the value of IRT including Rasch analysis and will act as a springboard for their use within the vision research community.

REFERENCES

1. Hays RD, Hahn H, Marshall G. Use of the SF-36 and other health-related quality of life measures to assess persons with disabilities. Arch Phys Med Rehabil 2002;83:S4–9.
2. Pesudovs K, Garamendi E, Elliott DB. The Quality of Life Impact of Refractive Correction (QIRC) Questionnaire: development and validation. Optom Vis Sci 2004;81:769–77.
3. Javitt JC, Steinberg EP, Sharkey P, Schein OD, Tielsch JM, Diener M, Legro M, Sommer A. Cataract surgery in one eye or both. A billion dollar per year issue. Ophthalmology 1995;102:1583–92.
4. Elliott DB, Patla AE, Furniss M, Adkin A. Improvements in clinical and functional vision and quality of life after second eye cataract surgery. Optom Vis Sci 2000;77:13–24.
5. Laidlaw DA, Harrad RA, Hopper CD, Whitaker A, Donovan JL, Brookes ST, Marsh GW, Peters TJ, Sparrow JM, Frankel SJ. Randomised trial of effectiveness of second eye cataract surgery. Lancet 1998;352:925–9.
6. Elliott DB, Trukolo-Ilic M, Strong JG, Pace R, Plotkin A, Bevers P. Demographic characteristics of the vision-disabled elderly. Invest Ophthalmol Vis Sci 1997;38:2566–75.

7. Lamoureux EL, Hooper CY, Lim L, Pallant JF, Hunt N, Keeffe JE, Guymer RH. Impact of cataract surgery on quality of life in patients with early age-related macular degeneration. Optom Vis Sci 2007;84:683–8.

8. Caudle LE, Williams KA, Pesudovs K. The Eye Sensation Scale: an ophthalmic pain severity measure. Optom Vis Sci 2007;84:752–62.

9. Hayes JR, Sheedy JE, Stelmack JA, Heaney CA. Computer use, symptoms, and quality of life. Optom Vis Sci 2007;84:738–44.

10. Borsting E, Chase CH, Ridder W, III. Measuring visual discomfort in college students. Optom Vis Sci 2007;84:745–51.

11. Kymes SM, Lee BS. Preference-based quality of life measures in people with visual impairment. Optom Vis Sci 2007;84:775–84.

12. Busbee BG, Brown MM, Brown GC, Sharma S. Cost-utility analysis of cataract surgery in the second eye. Ophthalmology 2003;110:2310–7.

13. Burr JM, Kilonzo M, Vale L, Ryan M. Developing a preference based glaucoma utility index using a discrete choice experiment. Optom Vis Sci 2007;84:797–808.

14. Likert RA. A technique for the measurement of attitudes. Arch Psychol 1932;140:1–55.

15. Jutai J, Day H, Woolrich W, Strong G. The predictability of retention and discontinuation of contact lenses. Optometry 2003;74:299–308.

16. Mangione CM, Phillips RS, Seddon JM, Lawrence MG, Cook EF, Dailey R, Goldman L. Development of the 'Activities of Daily Vision Scale'. A measure of visual functional status. Med Care 1992;30:1111–26.

17. Pesudovs K, Garamendi E, Keeves JP, Elliott DB. The Activities of Daily Vision Scale for cataract surgery outcomes: re-evaluating validity with Rasch analysis. Invest Ophthalmol Vis Sci 2003;44:2892–9.

18. Mallinson T. Why measurement matters when measuring patient visual outcome. Optom Vis Sci 2007;84:675–82.

19. Pesudovs K, Burr JM, Harley C, Elliott DB. The development, assessment and selection of questionnaires. Optom Vis Sci 2007;84:663–74.

20. Bond TG, Fox CM. Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, NJ: L. Earlbaum, 2001.

21. Massof RW. The measurement of vision disability. Optom Vis Sci 2002;79:516–52.

22. van Nipsen RM, Knol DL, Langelaan M, de Boer MR, Terwee CB, van Rens GH. Applying multilevel item response theory to vision related quality of life in Dutch visually impaired elderly. Optom Vis Sci 2007;84:710–20.

23. McKnight PE, Babcock-Parziale J. Respondent impact on functional ability outcome measures in vision rehabilitation. Optom Vis Sci 2007;84:721–8.

24. Massof RW. An interval-scaled scoring algorithm for visual function questionnaires. Optom Vis Sci 2007;84:689–704.

25. Lamoureux EL, Ferraro JG, Pallant JF, Pesudovs K, Rees G, Keeffe JE. Are standard instruments valid for the assessment of quality of life and symptoms in glaucoma? Optom Vis Sci 2007;84:789–96.

26. Langelaan M, van Nispen RMA, Knol DL, Moll AC, de Boer MR, Wouters B, van Rens GH. Visual Functioning Questionnaire: re-evaluation of psychometric properties for a group of working age adults. Optom Vis Sci 2007;84:775–84.

27. Court HJ, Greenland K, Margrain T. Content development of the Optometric Patient Anxiety Scale. Optom Vis Sci 2007;84:729–37.

28. Stelmack JA, Massof RW. Using the VA LV VFQ-48 in low vision rehabilitation. Optom Vis Sci 2007;84:705–9.

29. Hahn EA, Cella D, Bode RK, Gershon R, Lai JS. Item banks and their potential applications to health status assessment in diverse populations. Med Care 2006;44:S189–97.

30. Cella D, Gershon R, Lai JS, Choi S. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. Qual Life Res 2007.

31. Haley SM, Andres PL, Coster WJ, Kosinski M, Ni P, Jette AM. Short-form activity measure for post-acute care. Arch Phys Med Rehabil 2004;85:649–60.

32. Fries JF, Bruce B, Cella D. The promise of PROMIS: using item response theory to improve assessment of patient-reported outcomes. Clin Exp Rheumatol 2005;23:S53–7.

33. Kopec JA, Sayre EC, Davis AM, Badley EM, Abrahamowicz M, Sherlock L, Williams JI, Anis AH, Esdaile JM. Assessment of health-related quality of life in arthritis: conceptualization and development of five item banks using item response theory. Health Qual Life Outcomes 2006;4:33.

34. Haley SM, Coster WJ, Andres PL, Kosinski M, Ni P. Score comparability of short forms and computerized adaptive testing: Simulation study with the activity measure for post-acute care. Arch Phys Med Rehabil 2004;85:661–6.

35. Massof RW, Ahmadian L, Grover LL, Deremeik JT, Goldstein JE, Rainey C, Epstein C, Barnett GD. The Activity Inventory (AI): An adaptive visual function questionnaire. Optom Vis Sci 2007;84:763–74.

David B. Elliott
Bradford, UK
Konrad Pesudovs
Adelaide, Australia
Trudy Mallinson
Chicago, Illinois

David B. Elliott

Trudy Mallinson

Konrad Pesudovs