Giovanni Apolone
Paola Mosconi

Health-related quality of life (HRQOL) and migraine

G. Apolone (✉) • P. Mosconi
Laboratory of Clinical Research in Oncology, Mario Negri Institute of Pharmacological Research,
Via Eritrea 62, I-20157 Milan, Italy
e-mail: Apolone@marionegri.it
Tel.: +39-02-39014515
Fax: +39-02-33200231

Abstract Although the genuine interest in measuring qualitative aspects of life has increased in the medical field together with our capability to transform patients' subjective reports, opinions and ratings in standardized pieces of information, the debate about the true objective to measure, the focus and the terminology is still lively. Many instruments, both generic and disease-tailored, have been developed, tested and used in headache and migraine patients. Advances in the field of health outcome assessment have made available a new generation of approaches and measures, both specific and generic, that are based on an approach called dynamic assessment. According to the developers, it is now possible to produce valid and reliable estimates of a person's perception of health by asking only those questions relevant to the individual respondent's state. Despite the large body of evidence about the intrinsic value of such new methodology, its "incremental" value when applied in research projects or in the context of programs of disease or outcome management is yet not clear. Thus, before these new methods are experimentally compared with more traditional approaches, in the general population and in headache and migraine patients, caution should be recommended and their use should be avoided outside formal research projects.

Key words Quality of life • Health-related quality of life • Migraine • Computer adaptive testing • SF-36 • QualyMetric

Introduction

Migraine is a chronic condition that imposes a substantial burden on patients, families, employers and society. The impact of this burden is usually measured using traditional clinical outcomes and indicators of direct or indirect costs of illness [1]. In the last few years, patient-based (subjective) outcome measures, often called quality of life measures, have been used as primary or secondary outcomes, either in observational or experimental clinical studies where the values of different interventions are evaluated on different types of diseases and conditions [2, 3], including migraine [4, 5].

Although the genuine interest in measuring qualitative aspects of life has increased in the medical field [6] together with our capability to transform patients' subjective reports, opinions and ratings into standardized pieces of information [7–9], the debate about the true objective to measure (health status, happiness or satisfaction with life), the focus (general or disease-tailored values), and the terminology (quality of life or health-related quality of life) is still lively [10–12].

Quality of life (QOL), health-related quality of life (HRQOL) and subjective health status (HS) are terms that are used interchangeably in the medical field, though certain conceptual and operational differences among them do exist

[13]. QOL is a complex, abstract, multidimensional concept that defines an individual's satisfaction or happiness with life in domains he or she considers important. Often also referred to as "life satisfaction" or "subjective well-being", it is the broadest of all the concepts we may be interested in, and health is only one of the several dimensions of life usually considered in the models and taxonomies proposed so far.

HRQOL reflects an attempt to restrict the complex concept of QOL to those aspects of life specifically related to a person's health that potentially respond to health care. Most definitions of HRQOL include the domains of physical, mental and social functioning and well-being, as well as general health perceptions.

HS and health perceptions, also referred to as "perceived health status" are objective reports and subjective evaluations made by a person on his or her health. The need to distinguish between objective degrees of health status (reports) and subjective perceptions of health (evaluations) relies on the fact that individuals perceive themselves as healthy or ill independently of biological and physiological signs and symptoms of disease. In other words, two people with the same health status may have different perceptions of health.

Among the several approaches used to measure the previously mentioned concepts, when the aim is to use HRQOL as an outcome in clinical studies (i.e. a change in the patient's health status that may be attributed to a specific intervention), most of measures available are based on the psychometric approach. All the most modern psychometric instruments are now based on a conceptualization of HRQOL that is multidimensional and for which the patient is the source of information.

Usually, HRQOL instruments are classified according to their focus as *disease-specific* (i.e. measuring health concepts tailored to a specific disease and treatment, relevant and sensitive only to the condition under evaluation) or *generic* (i.e. assessing health concepts that represent basic human values, relevant to everyone's health). Examples of the first categories are the Short Form 36-item Health Survey and its shorter version SF-12 [14–17], available in Italian since 1995 and recently revised in a second version [18]. In addition, NHP [19] and SIP [20] are also available in Italian.

Many specific instruments have been developed, tested and used in headache and migraine patients. Examples are the Migraine Specific Questionnaire [21], MIDAS [22], MS-QoL [23] and HIT-6 [24].

Advances in health outcome assessment

Recently, advances in the fields of health outcome assessment have made available a new generation of approaches and measures, both specific or generic, that are based on a novel approach called dynamic assessment. From a pool of items from established health surveys, using computerized algorithms, a brief but valid assessment of concepts related to health and healthcare are administered to subjects by asking only those questions relevant to the individual under evaluation. On average, such dynamic and customized assessment is shorter than the corresponding fixed forms and requires less time for completion. On the other hand, it is supposed to have the same precision and interpretability and, when delivered as a Web-based survey, it allows immediate feedback reporting.

Such innovation enables us to classify all the HRQOL measures in 2 new distinct categories: fixed length (having a fixed number of questions) such as all the so far available questionnaires, and dynamic length (having a variable number of questions tailored to the kind of answers given by interviewed patients).

Objective of the present paper is to introduce the concept and the measures pertaining to the field of dynamic assessment and to discuss the pros and cons of "old" and "novel" measures in the context of the present scenario. Examples are taken from 2 Web sites developed and implemented by QualyMetric, a US company, that has published online demonstrations of its proprietary tools. Utilization of these demonstrative tools implies that users are aware of the terms of use and of the usage rules that are well described in the relevant sites (www.qmetric.com and www.amIhealthy.com).

Computer adaptive testing

Modern psychometric methods, such as item response theory (IRT) and Rasch analysis, together with a wider availability of computer and Web technologies have made possible an easier and more friendly implementation of dynamic (computer-assisted) assessment approach, known for a long time and recently applied to headache and migraine [24]. IRT and Rasch analysis are statistical models for measuring the relationship between a person's response to an item and his or her score on a concept or latent trait. In contrast to traditional psychometric methods, they do not assume that measurement error is constant across the various levels of a given scale, but assume that a specific estimate of the measurement error can be given for each person at each scale level. Accordingly, measurement error is determined by the constellation of the item characteristics, such as difficulty and discrimination parameters answered by the respondent. Therefore, the instrument-independent characteristics of Rasch mean that a respondent's ability for a given construct is assessed independently of the particular characteristics of the administered items. Examples of the application of Rasch and IRT methods have been published elsewhere, for the SFs [25] and for migraine-specific measures [24].

Thus, it is possible to derive comparable scores on a particular construct for respondents who have completed different items, pertaining to the same construct. In addition, once a model has been estimated for a pool of items, the model can be used to select items for a specific purpose. The availability of a large database containing several items and computer algorithms makes possible the identification of the optimal selection of items for a given individual, resulting in a short list of questions and a more precise estimate of his or her health measure.

As described in detail in the Web site of QualyMetric, the availability of a large pool of items from well established health surveys (either general and disease-specific) and the application of Rasch and IRT approaches allow for assessments to be individually tailored or adapted for test responders. In this way, it is now possible to produce valid and reliable estimates of the subject perception of health by asking only those questions relevant to the individual respondent's state. In addition, by scoring all responses on a standard metric, results can be compared for those who answer different questions. According to developers, the new measures are delivered in a fraction of the time and cost of traditional (fixed-length) assessment. In addition, the new tools enhance the possibility of using patients' reported data for several purposes, such as monitoring health, predicting outcomes and screening diseases.

At present, several (proprietary) tools are available as Internet applications. Demos of products assessing general aspects of health (DyHA General Health) or headache (DYHA Headache Impact Survey) are available in the QualyMetric Web site and in other sites. After registering, it is possible to perform a free test of some fixed and dynamic approaches. At the moment, version 2 of the SF-36 and the dynamic general health assessment (that basically is based on the pool of items from which the SF-36 was statistically derived in the past) are available to test the performance of both methods.

The final output is quite similar in terms of point estimates, level of information and interpretability. What is dramatically different is the number of questions actually delivered and the time required for the completion

A similar exercise (e.g. comparing the fixed and dynamic approaches) may be done on the same Web sites using two tools addressing headache-tailored questions, the HIT-6 and the Headache Impact Test questionnaires.

Conclusions

Despite the lively debate still ongoing about conceptual and operational issues, interest in measuring relevant qualitative aspects of life that are most closely related to health and health-care has grown in recent years.

In addition to the genuine goal to increase patients' involvement in medical decision-making through the use of measures based on patients' perception (thereby introducing the patients' point of view into clinical research), not infrequently HRQOL measures are included as a therapeutic efficacy endpoint in several industry studies, both to assist the industry in the regulatory process and for marketing purposes.

Nevertheless, regulatory agencies such as FDA and EMEA currently do not require this kind of data. A recent review of all the official documents produced and circulated by EMEA containing recommendations on trial design, conduct and analysis documented that none of the 189 documents retrieved and evaluated focused directly on HRQOL. Only few explicit recommendations were identified for 13 specific drugs or conditions, and most of the time (12 out of 13) HRQOL was recommended as a secondary or complementary endpoint [26].

Measurement of clinical and subjective health outcomes has become increasingly important in our rapidly changing healthcare environment where the objectives to minimize costs and maximize outcomes have stimulated great interest in methods to assess and monitor of the effects of healthcare interventions on clinical, humanistic and economic outcomes. The hub of these new efforts is the availability of short, valid, friendly and inexpensive measures to determine the value of healthcare or medical interventions delivered to the population. Recent advances, such as the combination of the new psychometric approach with dynamic (computer-assisted) assessment implemented through Internet and the possibility to have real-time results, have the potential to increase the use of HRQOL measures outside the domain of research. In fact, several companies are now marketing these new approaches claiming their multipurpose nature and thus suggesting their use in the context of programs of disease or outcome management to screen for subjects with the greatest risk of poor health outcome or the highest probability to consume health resources. There are clear dangers involved, as the field of health outcome assessment and outcome research has already witnessed negative effects with the widespread adoption of tools that have ended up being both appropriately and inappropriately used [27, 28].

Although a large body of evidence does exist about the validity and reliability of "old" and "novel" subjective measures when applied in research projects, less evidence is available about their actual value when implemented outside the research setting. The main questions regard *whether* we need these subjective measures to assess the efficacy, quality and costs of medical interventions. In addition, scarce evidence is available about the incremental validity of the new and more sophisticated approaches presented and discussed herein.

Before implementing a new instrument, tool or technology in clinical practice, its validity must be tested in terms of effectiveness, safety and cost. The cornerstone of this

evaluation is a formal comparison of the new technology with an appropriate standard in representative samples. New HRQOL measures should be tested in this way. Migraine is a suitable condition where such comparison can be performed as generic and specific measures, both fixed and dynamic, are already available. Although dynamic approaches are promising, before the results of such confirmatory efforts are available, caution should be recommended and their use should be avoided outside formal research projects.

References

1. Solomon GD, Price KL (1997) Burden of migraine: A review of its socioeconomic impact. Pharmacoeconomics 11(2):229–237
2. Testa MA, Simonson DC (1996) Assessment of quality of life outcomes. N Engl J Med 334:835–840
3. Guyatt GH, Feeny DH, Patrick DL (1993) Measuring health related quality of life. Ann Int Med 118:622–629
4. Solomon GD (1997) Evolution of the measurement of quality of life in migraine. Neurology 48(3):S10–S15
5. Liddel J (1994) Migraine: the patient's perspective. Contemp Pharmacother 5:253–257
6. Spilker B, Revicki DA (1995) Taxonomy of quality of life. In: Spilker B (ed) Quality of life and pharmacoeconomics in clinical trials, 2nd edn. Lippincott Raven, Philadelphia, pp 25–31
7. Fitzpatrick R, Fletcher AE, Gore SM, Jones DR, Spiegelhalter DJ, Cox DR (1992) Quality of life measures in health care. I: Applications and issues on measurement. BMJ 305:1074–1077
8. Stewart AL, Ware JE (1992) Measuring functioning and well-being. The Medical Outcomes Study approach. Duke University, Durham London
9. Tarlov AR, Ware JE, Greenfield S, Nelson EC, Perrin E, Zubkoff M (1989) The Medical Outcome Study. An application of methods for monitoring the results of medical care. JAMA 262:925–930
10. Leplège A, Hunt S (1997) The problem of quality of life in medicine. JAMA 278(1):47–50
11. Apolone G (1998) Defining and measuring quality of life in medicine. JAMA 279:431
12. Patrick DL, Deyo RA (1989) Generic and disease-specific measures in assessing health status and quality of life. Med Care 27(3):S217–S232
13. Apolone G, Mosconi P (1998) Review of the concept of quality of life assessment and discussion of the present trend in clinical research. Nephrol Dial Transplant 13:65–69
14. Ware JE, Snow KK, Kosinski M, Gandek B (1993) SF-36 Health Survey. Manual and interpretation guide. The Health Institute, New England Medical Center, Boston
15. Apolone G, Mosconi P, Ware JE (1997) Questionario sullo stato di salute SF-36: Manuale d'uso e guida all'interpretazione dei risultati. Guerini ed Associati, Milan
16. Ware JE, Kosinski M, Keller SD (1995) SF-12: How to score the SF-12 physical and mental health summary scales. The Health Institute, NEMC, Boston
17. Apolone G, Mosconi P, Quattrociocchi L, Gianicolo EAL, Groth N, Ware JE (2001) Questionario sullo stato di salute SF-12, versione italiana. Guerini ed Associati, Milan
18. Ware JE, Kosinski M, Dewey JE (2000) How to score version 2 of the SF-36 health survey. Quality Metric, Boston
19. Hunt SM, McEwen J, McKenna SP (1986) Measuring health status. Croom Helm, London
20. Bergner M, Bobbit RA, Carter WB, Gilson BS (1981) The sickness impact profile: development of and final revision of a health status measures. Med Care 19:797–805
21. Jhingran P, Osterhaus JT, Miller DW, Lee JT, Kirschdoerfer L (1998) Development and validation of the migraine-specific quality of life questionnaire. Haedache 38:295–302
22. Stewart WF, Lipton RB, Whyte J, Dowson A, Kolodner K, Liberman JN, Sawyer J (1999) An international study to assess reliability of the migraine disability assessment (MIDAS) score. Neurology 53:988–994
23. Wagner TH, Patrick DL, Galer BS et al (1996) A new instrument to assess the long-term quality of life effects from migraine: development and psychometric testing of the MSQoL. Headache 36:484–492
24. Ware JE, Bjorner JB, Kosinski M (2000) Practical implications of item response theory (IRT) and computer adaptive testing: a brief summary of ongoing studies of widely used headache impact scales. Med Care 38(II):S73–S82
25. Raczek AE, Ware JE, Bjorner JK, Gandek B et al (1998) Comparison of Rasch and summated rating scales constructed from SF-36 physical functioning items in seven countries. Results from the IQOLA project. J Clin Epidemiol 51:1203–1214
26. Apolone G, De Carli G, Brunetti M, Garattini S (2001) HR-QOL and regulatory issues. An assesssmnet of the European Agency for the Evaluation of Medicinal Products (EMEA) recommendations on the use of HR-Qol measures in drug approval. Pharmacoeconomics 187–195
27. Epstein RS, Sherwood LM (1996) From outcomes research to disease management: a guide for perplexed. Ann Intern Med 124:832–837
28. Harris JM Jr (1996) Disease management: new wine in new bottles? Ann Intern Med 124:838–842