



Value Insider Season I Episode 2: How to Measure Quality of Life and Utility? (QoL) [Podcast]

Nancy J Devlin ¹, Elisabeth Sophia Hartgers-Gubbels ², Michael Chambers³

¹Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia; ²Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany; ³MC Healthcare Evaluation, London, UK

Correspondence: Elisabeth Sophia Hartgers-Gubbels, Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany, Email lisa.hartgers-gubbels@boehringer-ingelheim.com

Abstract: How do we measure quality of life and utility of interventions? In this episode of the Value Insider podcast, host Mike Chambers speaks with Prof. Nancy Devlin about incorporating quality of life in value assessment. Prof. Devlin is professor of health economics at the University of Melbourne. Her past roles include Director of Research at the Office of Health Economics (OHE) London, Professor of Health Economics at City University of London, and she has been director of the International Society of Pharmacoeconomics and Outcomes Research (ISPOR). She is the Chair of the Board of the EuroQol Research Foundation, the international research organization which has developed and maintains the EQ-5D patient-reported outcome (PRO) instrument. Prof Devlin explains the value of the patient voice and how it can be measured and taken into account when considering the value of healthcare interventions.

Keywords: market access, healthcare reimbursement, health technology assessment, HTA, value demonstration, health economics and outcomes research, payer

Host: Mr Mike Chambers (MC), Independent expert

Guest: Prof. Nancy Devlin (ND), Independent expert

This [podcast](#) is brought to you by Boehringer Ingelheim

Chapter I: General Introduction [00.00]

MC: Welcome to the Value Insider podcast series. In this series, with the help of experts in the field, we will be exploring the fundamentals of assessing value in healthcare, especially when looking at the value of new healthcare interventions.

My name is Mike Chambers, I am founder and director of MC Healthcare Evaluation, and I have spent the last twenty-five years working in health economics and health technology assessment for the pharmaceutical and medical diagnostics industries, and more recently as an independent advisor. I am also a member of the Technology Appraisal Committee at NICE: the National Institute for Health and Care Excellence in the UK. It is my great pleasure to be your moderating host for this season of the Value Insider podcast series.

Chapter 2: Episode Introduction and Welcome [00.58]

MC: Thank you for joining us today, be sure to subscribe and follow this series to ensure that you do not miss any of the informative podcasts in the series.

Today, we will be talking about the value of the patient voice and how it can be measured, with our guest speaker Professor Nancy Devlin. Nancy is professor of health economics at the University of Melbourne. Her past roles include Director of Research at the Office of Health Economics, and Professor of Health Economics at City University, both in London, and she has been director of the International Society of Pharmacoeconomics and

Outcomes Research. She is the Chair of the Board of the EuroQol Research Foundation, the international research organization which has developed and maintains the EQ-5D instrument. Welcome Nancy.

ND: Thank you, Mike, it's great to be here.

Chapter 3: The Patient Voice [01.44]

MC: So, let us start. Why is the patient voice important in healthcare?

ND: Well, because the role of most healthcare is to improve how patients feel. And the patients themselves are the best judges of how they feel. They are the experts on their own health. So a starting point for any health technology appraisal, is going to be the patient voice.¹⁻³ Particularly for things that are subjectively experienced by the patients, such as pain, fatigue, emotional distress, their satisfaction with care, the impact of their symptoms on their daily life, their ability to do their day-to-day activities, and so forth. Measuring the extent of which an intervention improves those kind of outcomes, is going to be really important to understanding whether a treatment is going to be effective, or not.

Chapter 4: Patient-Reported Outcomes (PROs) [02.33]

MC: Health-related quality of life is often measured using patient-reported outcomes, or PROs - what are PROs?

ND: Patient-reported outcomes, they are a systematic, structured way of capturing the patients' feelings about their own health,⁴ which is subjective, but to do that in a systematic and robust way,⁵ that enables us to compare those measurements across patients.

MC: PROs are used to collect information about health-related quality of life, but also symptom scores, and maybe patient experience.

ND: That's an interesting point. There are a separate category of things called PREMs, or patient-reported experience measures, that aim to capture the patients' experience of their treatment, or of the care or healthcare services that they are receiving. Sometimes, for example, their satisfaction of the care that they are receiving.^{6,7} Whereas patient-reported outcome measures are really capturing the patients' views and feelings about their health.

MC: We are talking about the use of patient-reported outcomes to measure health-related quality of life. Is there a difference between health-related quality of life and quality of life?

ND: Ah, that's a very tricky question. [laughs] Some people describe patient-reported outcomes as measuring health status. A few decades ago, that's what we would have all said. That these are questionnaires, that measure the patients' health status. Increasingly they are now described as measuring health-related quality of life. The impacts of health on those aspects of their functioning, and day-to-day activities. Things that patients experience from their health like whether they have got pain, whether they are feeling depressed and anxious, and so forth. But there is a dividing line between that and what we might describe as quality of life more generally.

Chapter 5: Disease-Specific and Generic PROs [04.33]

MC: When measuring health-related quality of life, using PROs, we can use disease-specific and generic instruments. And by instruments I think we generally mean questionnaires in this context. Could you explain the difference between generic and disease-specific instruments?

ND: Certainly. And um, they sort of do what they say on the tin, really. A disease-specific questionnaire is a questionnaire which has all of the questions in it aiming to measure health related to a specific condition, or sometimes a specific part of the body. So the Oxford Hip score,⁸ for example, is capturing aspects of mobility, discomfort, mobility that relate to problems with hips and joints. And often these are somewhat longer questionnaires, with multiple questions in them, providing a very rich and detailed account of the kind of problems that patients experience, with specific conditions.

The generic questionnaires, on the other hand, again the hint is in the name. They aim to ask questions that let us measure health in a general way, it's not specific to any disease or illness, but rather tries to capture the main aspects of health that really matter to all patients. And the idea with the generic instruments is that we can measure health in a manner that we can compare that measure between totally different kinds of patients. Of different ages, with different conditions.

MC: And do you have a good example to share of a generic instrument?

ND: Yes! The SF36 is a 36-question instrument,⁹ very, very widely used. Breaks down into a mental component score, and a physical component score. So there's not an overall score that comes from it, but rather two scores that capture different aspects of the patients' self-reported health. There are much shorter questionnaires of a generic type available. There's the HUI,¹⁰ that was developed in North America, Health Utility Index. The most widely used one in health technology appraisal, is the EQ-5D, which was developed by the EuroQol Group.¹¹ It's just five questions, just broken down most of the things that might matter to most patients.

Chapter 6: Developing a PRO [07.08]

MC: How easy is to develop a new patient-reported outcome instrument?

ND: It's a really long and arduous process. It's a process that can take literally years from conception through to initial development, through to validity testing, through to piloting, and then revisiting.¹² Normally such instruments, as they are developed in a good manner, are developed to have what we call good "psychometric properties", and that means that they are measuring what they are set out to measure, they are doing that in a reliable way, and that when the person's underlying health changes, that we see the changes that we might expect to see in the measures that we are capturing. All of that can literally take years. For that reason I tend to discourage people from developing their own new instruments. There are literally thousands of patient-reported outcome instruments out there.

MC: A task not to be taken lightly.

Chapter 7: Choosing a PRO [08.11]

MC: Out of those thousands of PROs that are available, many of which will not have been developed, historically, in the way you describe. How do you choose which ones to use in your study?

ND: Well that's a great question, because there are so many of these instruments around. Key thing is to begin with a literature review, and to check what the evidence is on their performance: how the instruments work to describe patient health. So obviously they should be relevant to the patient's experience and the patient pathway. So, does the questionnaire measure the things that patients find really important, is there good evidence to support that? And then we will be looking for evidence on the measurement properties of the instruments.

And there should be evidence arising from studies that have been specifically undertaken to test these things. So, are the questionnaires reliable? Are you able to repeat the questionnaires with similar patients and obtain similar outcomes? Are they measuring what they are intending to measure? And do they have validity in terms of the constructs, or the nature of what's actually being measured. If the patients' health changes over time, are the measures detecting that.

A related point, which is somewhat more complex and subtle, is that you might also want to take into account what instruments are popular with clinicians in that area. Now if you are extremely lucky, the instruments which are widely used and popular with clinicians, will also be the ones with the best psychometric properties. But it's not always the case. It can be very difficult to displace a well-loved and well-used instrument.

Chapter 8: The Wider Stakeholder Environment [10.01]

MC: My understanding also is that regulatory agencies, such as FDA and EMA have strict requirements about PROs, before they are considered to be acceptable. Is this an important consideration?

ND: It's a very important consideration. Regulatory agencies like FDA,¹³ EMA,¹⁴ generally have quite strict requirements about the nature of the patient-reported outcome questionnaires, that they accept as legitimate evidence. And that means that they can be allowed to be used in the marketing of new medicines.

MC: Are there differences between the use of PROs in clinical studies, and trials especially, and their use in observational studies, so-called Real-World studies, or routine clinical practice for that matter?

ND: These patient-reported outcome measures are used in all of those things. The EQ-5D,¹¹ for example, is used in clinical trials, and it has a very specific purpose for being used in clinical trials, It's also used in routine, day-to-day, clinical settings. There are also observational studies where we might be conducting a study with a sample, but where they are receiving care in a real-life setting. And then there's also population health surveys and those

are important too, because they give you a kind of benchmark of what the healthy population reports in terms of health.

MC: I am aware of some shorter instruments that are used in routine practice. Not just for research, but for clinical decision-making and to inform conversations between clinicians and patients. I guess they have to be shorter and simpler to allow that conversation to occur.

ND: Yes, that's correct, and I guess the other thing about routine outcomes measurements in clinical settings, is, time is short, both for the patient and for the clinician. And also the volumes of data which have been collected in routine outcomes measurement, is absolutely huge. There's also thinking about, you know, how will those data be used, what kind of analyses do we want to perform on them? And most importantly of all, what decisions do we think the evidence from that data will inform? So, you know, it should always have a very clear aim.

Chapter 9: Quality-Adjusted Life Year (QALY) [12.18]

MC: Let us move on to valuation of quality of life. So, quality of life is often valued in terms of something called a QALY, a Quality-Adjusted Life Year.

ND: Well, the QALY¹⁵ combines two things: the length of life of patients, and the quality of their life in survival. So the quality of life is considered very broadly and includes things like mobility, the patients' ability to care for themselves, their ability to do their daily activities, their pain, anxiety, and so forth. And the QALY provides a kind of common currency for allowing comparisons of health improvements to be made across very different kinds of diseases and health problems. It provides a kind of common metric.

MC: And how are QALYs calculated? Can you just give an insight into that?

ND: The starting point is the patients' measure of their own health, on patient-reported outcomes. Once the patient has completed those questionnaires, the data are summarized by values on a zero-to-one scale that tell us the quality weight, how good or bad those health states are, overall. That information is then combined with length of life, to give us a quality-adjusted life year.

MC: And the strength of the QALY is that it can value quality of life across a whole range of disease areas, which decision-makers may be very interested in being able to compare across.

ND: That's exactly right. Within health technology appraisal, the key is to be able to have a measure of outcomes, so the QALY, which can be compared across totally different kinds of conditions. That's really important, because decision-makers are trying to weigh up which treatments are going to yield the most value for patients, and budgets are not unlimited.

And to understand where these quality of life values or weights come from, members of the general public are selected to take part in a survey. And we ask them all sorts of different questions and tasks, to help us understand what types of health problems matter more to people, compared to others. And that's one thing that we know with absolute certainty, that different types of health problems mean more to some people than others. They are more detrimental, considered more problematic. And that's really important, because it means that treatments that improve those problems, should be given a much higher value when we evaluate them.

And that is a really important difference with the condition-specific instruments we described earlier on, with a kind of arbitrary scoring system, that is basically added up in a linear way. And that assumes that all of the items in that questionnaire carry exactly the same weight. They are all equally important. And that's fine when you are just trying to describe a condition, but actually if you are trying to really reflect how good or bad a health problem is, you really need that preference data to come into play.

MC: Thank you. Not all countries accept QALYs, or find them in favor. I am thinking of Germany, and to a large extent the United States. Why is this?

ND: Germany considers the effectiveness and cost-effectiveness of new technologies, using a focus which is very much on the disease-specific level. QALYs are not generally required in Germany because the decision-makers there pay less attention to achieving value for money right across the entire budget, and do not require comparisons between disease areas.

US payers, different again, they tend to focus on relative or comparative effectiveness rather than cost-effectiveness, as well as on budget impact or affordability.

MC: Please listen to our next episode if you'd like to hear more about that.

Are there any alternatives to QALYs?

ND: Certainly in the field of lower- and middle-income countries. Instead of using QALYs as a measure of outcome and improved health from care, they tend to do things in a manner which is almost the mirror image of it, which is using disability-adjusted life years.¹⁶ And the focus of disability-adjusted life years is to measure the burden of

disease. So DALYs are measuring how much health you have not got. It's a measure which was developed and is favored by the World Health Organization. They measure years of healthy life lost due to premature mortality and disability. They do not fully capture quality of life, and they are not often used as a means of evaluating specific technologies. For that purpose, QALYs are much more often used and regarded as the standard, for health technology appraisal.

Chapter 10: PROs and QALYs in HTA [17.28]

MC: How are PROs and QALYs used in Health Technology Assessment?

ND: They form a really important way in which evidence from patients feeds into decision-making.² So the starting point is always the clinical trials, and embedding the patient-reported outcomes into those trials, at baseline for example, we take a measure of how patients feel, and by observing what's is happening to the patients in the active treatment arm, against placebo or against other treatments, we can observe the extent to which their health improves, from the point of view of the patient. In other words, what do we know about this treatment, and the extent to which it extends life and improves quality of life.

Then, we add to that, the fact that treatments are not free, of course, so we want to know: okay, this treatment has improved the patients' survival, it has improved the patients' quality of life, but it also costs a little bit more than existing treatments. All of that information is then combined, by looking at the addition to cost per QALY gained. And you can see there is an awful lot wrapped up in that equation. They can be quite complex models, with lots and lots of different arms and decision trees and probability weightings.

MC: We will hear more about that sort of thing in one of our future episodes.

My experience of working in the UK in health technology assessment is that this QALY measure is very attractive in that it is a single measure, which can help with decision-making. But there's a huge richness of PRO, quality of life data that are presented to the agencies, which we do not do as well as we can, I think, in terms of drawing from.

ND: I am really glad you raised the point, because it's something I emphasize over and over again, to people who collect patient-reported outcome data from patients. Look at the data that the patients give you! I think there is a real tendency, especially where the aim is to submit to HTA bodies, to just immediately summarize that data by these zero-to-one values. And that's fine, but you lose a lot of information, when you do that. I would always recommend looking at the individual questions that you have asked, and the patients' responses to those. There are lots of really easy ways of analyzing that data, that can tell you an awful lot about how the patients feeling, and exactly how their problems are alleviated. And also, potentially, made worse by side effects. And unless you really drill down into that level of detail of what the patients have told you, you might miss an awful lot of what's important.

Chapter 11: EQ-5D and Alternatives [20.18]

MC: You have been centrally involved in the development of this special questionnaire which is called the EQ-5D, which is specifically designed to generate QALYs. What is important about the EQ-5D?

ND: Its strengths are that it's concise, so it only has five questions in it. Now, you know, obviously you are not going to pick up absolutely everything that matters to all patients. But it is surprisingly powerful in picking up what matters to most patients, most of the time. The fact that it's concise, means that you can readily add that in to both a large population survey, you can ask it on a mobile phone using digital devices, and it's not going to be overly burdensome on the respondent. So, it's concise, and it's super generic. Enables to compare the extent to which these, at the end of the day, really end up improving patients' quality of life.¹¹

MC: There are other instruments that are used. You have already mentioned the HUI, and I know there is a way of generating QALYs from the SF36.

ND: That's right. And I think the advantage that seem to be associated with that, is the confidence that you are using the same measure each time, so the estimates of QALYs mean the same thing. So a kind of slogan for health technology appraisal, is that "a QALY, is a QALY, is a QALY", meaning that if you get a QALY estimate from one disease area, you can safely assume that it has the same measurement properties to evidence on QALYs submitted from a totally different disease area.

The disease-specific information is really important as well, but it does not give you that comparability. And for an organization like NICE,¹⁷ which is constantly receiving evidence across totally different conditions, they

want to be confident that whatever estimate of QALY or other outcomes they are using, that they are making broadly coherent decisions about value-for-money, right across all of the different technologies submitted to it.

Chapter 12: Pediatric Population [22.26]

MC: And I know some of your recent research has focused on the development of patient-reported outcomes, the measurement of quality of life, in children. Pediatric populations. Maybe even the generation of QALYs for children. So, could you tell us a little bit about that work?

ND: Delighted to. It's one of the most challenging areas, I think, for health economists. And we have done a number of reviews of decision-making by both NICE¹⁷ and by PBAC¹⁸ and elsewhere, to look at the quality of the evidence submitted around technologies aimed at children. And the conclusion from those reviews is always the same: the standard of the evidence is appallingly bad, overall. Huge gaps in evidence. So the HTA bodies around the world recognize this is a huge gap, and there are lots of technologies coming through the pipeline, which actually are aimed at children. Particularly, you know, genetic testing, lots of orphan diseases where treatments are developed.

So in our research program,¹⁹ we have been developing new measures. One of the particular challenges here, which is an obvious one, is that children cannot self-report their own health, below a certain age. And so we rely on proxy-completion. But you have to be really confident that those completing instruments on behalf of children, are truly representing the children's perspective. We are developing animation and pictographs so that children of younger ages can report their own health. We are seeking values from the general public for quality of life in children, so that we are not just using adult values for that. That's tended to be what's happened in the past. And we are also looking at whether QALYs are given higher priority in children, compared to adult QALYs. So there's a huge black hole in the evidence, I am afraid, around child QALYs at present, but we are doing an awful lot of work in making progress quite quickly.

Chapter 13: Conclusion [24.28]

MD: Nancy thank you for that, and we look forward to the outputs of your research to help fill that black hole. Thank you, everybody, for joining us today in this podcast. I would like especially to thank Professor Nancy Devlin for an engaging conversation and for giving us a comprehensive overview of the value of the patient voice in healthcare.

I hope you can join us for the third podcast in this series with Professor Sean Sullivan, at the University of Washington, Seattle, where we will focus on how healthcare payers see value, in particular in terms of affordability and budget impact. If you have enjoyed this podcast, please subscribe to our series, and thank you for listening.

Acknowledgment

The authors thank Rachel Emerson-Stadler (Boehringer Ingelheim International GmbH, Germany) for her early contributions in the design of the podcast series and most particularly Dr. Lindsay Nicholson (Maverex Ltd, UK) for her expert assistance at all stages in the development of this episode, including diligent background research and curation of content under direction of the authors.

Disclosure

Those involved have been compensated for their time by Boehringer Ingelheim. Speakers on this podcast speak on personal behalf, independently of Boehringer Ingelheim. E.S. Hartgers-Gubbels is an employee of Boehringer Ingelheim International GmbH. M Chambers is the founder of MC Healthcare Evaluation and a former employee of GSK and GE Healthcare. N.J Devlin is a member of the Euroqol Group. Views expressed in this work do not necessarily represent the views of the EuroQol Research Foundation. The authors report no other conflicts of interest in this work.

References

1. Perfetto EM, Oehrlein EM, Boutin M, Reid S, Gascho E. Value to whom? The patient voice in the value discussion. *Value Health*. 2017;20(2):286–291. doi:10.1016/j.jval.2016.11.014
2. Mott DJ. Incorporating quantitative patient preference data into healthcare decision making processes: is HTA falling behind? *Patient*. 2018;11(3):249–252. doi:10.1007/s40271-018-0305-9
3. Johnson DS, Bush MT, Brandzel S, Wernli KJ. The patient voice in research—evolution of a role. *Res Involv Engagem*. 2016;2(1):1–6. doi:10.1186/s40900-016-0020-4
4. Deshpande PR, Rajan S, Sudeepthi BL, Nazir CA. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res*. 2011;2(4):137. doi:10.4103/2229-3485.86879
5. Prinsen CA, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res*. 2018;27(5):1147–1157. doi:10.1007/s11136-018-1798-3
6. Basch E, Stover AM, Schrag D, et al. Clinical utility and user perceptions of a digital system for electronic patient-reported symptom monitoring during routine cancer care: findings from the PRO-TECT trial. *JCO Clin Cancer Inform*. 2020;(4):947–957. doi:10.1200/cci.20.00081
7. Nielsen AS, Appel CW, Larsen BF, et al. Patient perspectives on digital patient reported outcomes in routine care of inflammatory bowel disease. *J Patient-Rep Outcomes*. 2021;5(1):1–12. doi:10.1186/S41687-021-00366-2/TABLES/2
8. Wylde V, Learmonth ID, Cavendish VJ. The Oxford Hip score: the patient's perspective. *Health Qual Life Outcomes*. 2005;3(1):1–8. doi:10.1186/1477-7525-3-66
9. Ware JE Jr. SF-36 health survey update. *Spine*. 2000;25(24):3130–3139. doi:10.1097/00007632-200012150-00008
10. Mo F, Choi BC, Li FC, Merrick J. Using Health Utility Index (HUI) for measuring the impact on health-related quality of life (HRQL) among individuals with chronic diseases. *TheScientificWorldJOURNAL*. 2004;4:746–757. doi:10.1100/tsw.2004.128
11. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. *Appl Health Econ Health Policy*. 2017;15(2):127–137. doi:10.1007/s40258-017-0310-5
12. Atkinson MJ, Lennox RD. Extending basic principles of measurement models to the design and validation of patient reported outcomes. *Health Qual Life Outcomes*. 2006;4(1):1–12. doi:10.1186/1477-7525-4-65
13. FDA. Patient-reported outcome measures: use in medical product development to support labeling claims - Guidance for industry; 2009. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims>. Accessed September 27, 2022.
14. Committee for Medicinal Products for Human Use. *Reflection Paper on the Regulatory Guidance for the Use of Health-Related Quality of Life (HRQL) Measures in the Evaluation of Medicinal Products*. London: European Medicines Agency; 2005.
15. Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to pay for a quality-adjusted life year: in search of a standard. *Med Decis Making*. 2000;20(3):332–342. doi:10.1177/0272989X0002000310
16. Anand S, Hanson K. Disability-adjusted life years: a critical review. *J Health Econ*. 1997;16(6):685–702. doi:10.1016/S0167-6296(97)00005-2
17. National Institute for Health and Care Excellence (NICE). About [Homepage]. Available from: <https://www.nice.org.uk/about>. Accessed September 01, 2022.
18. Pharmaceutical Benefits Advisory Committee (PBAC). Pharmaceutical Benefits Advisory Committee (PBAC) Membership. Available from: <https://www.pbs.gov.au/pbs/industry/listing/participants/pbac>. Accessed September 01, 2022.
19. Quality of Life in Kids: Key evidence to strengthen decisions in Australia (QUOKKA) Research Program. Available from: <https://www.quokkaresearchprogram.org/>. Accessed September 01, 2022.

International Journal of General Medicine

Dovepress

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>