Use of surrogate endpoints in healthcare policy: proposal for consistent adoption of a validation framework

Oriana Ciani, Marc Buyse, Mike Drummond, Guido Rasi, Everardo D. Saad and Rod S. Taylor

Evidence Synthesis and Modelling for Health Improvement, Institute of Health Research, University of Exeter Medical School, Exeter, UK (Oriana Ciani, PhD, Rod S Taylor, PhD); Centre for Research on Health and Social Care Management, Bocconi University, Milan, Italy (Oriana Ciani, PhD); International Drug Development Institute, Louvain-la-Neuve, Belgium (Marc Buyse, ScD); Centre for Health Economics, University of York, York, UK (Mike Drummond, PhD); European Medicine Agency, London, UK (Guido Rasi, PhD); University "Tor Vergata", Rome, Italy (Guido Rasi, PhD); Dendrix Research, Sao Paulo, Brazil (Everardo D. Saad, MD)

Correspondence to: Dr Oriana Ciani, Institute of Health Research, University of Exeter Medical School, South Cloisters, St Luke's Campus, Exeter EX1 2LU, UK
o.ciani@exeter.ac.uk

Market access and coverage policies for health technologies should ideally be based on clinical trials that assess final outcomes relevant to patients, such as survival, morbidity, and health-related quality of life. Nevertheless, growing pressure for faster access to innovative treatments for patients in the past two decades has led to the introduction of various regulatory initiatives intended to facilitate this[1]. Consequently, regulatory authorities and payers often have to base their decisions about the use of a technology on surrogate outcomes, which allow trials to be done less expensively with fewer patients in a relatively short period. For example, more than 40% of trials used as the basis for approval of new indications by the US Food and Drug Administration (FDA) between 2005 and 2012 had a primary outcome that was a surrogate endpoint[2]. Furthermore, evidence from surrogate endpoints may not only expedite the regulatory approval of new health technologies but also inform coverage and reimbursement

decisions. Over the last decade, between 27 to 50% of submissions to the National Institute for Health and Care Excellence (NICE) in the United Kingdom, the Pharmaceutical Benefits and Advisory Committee in Australia, and the Common Drug Review in Canada were based on surrogate endpoints[3].

However, relying on surrogate endpoints can pose several serious problems for healthcare decision-makers. First, surrogate endpoints may not capture the combined risk–benefit profile of a health technology[4]. Some drugs approved on the basis of surrogate endpoints have subsequently been associated with serious safety problems and have had to be withdrawn from the market or have their indications substantially restricted. Even if the surrogate lies in the only causal pathway of the disease process, reliance on surrogate endpoints can often lead to overestimation of the magnitude of the treatment effect on the final endpoint[5].

Second, the use of surrogate endpoint data requires decision-makers to extrapolate beyond the observed findings in order to estimate the expected true benefits to patients and health systems. Thus, clinical superiority on a surrogate endpoint may not necessarily translate into benefits that are good value for money for healthcare systems. This can be illustrated by the example of dasatinib, which has been approved by the European Medicines Agency for the treatment of chronic myeloid leukaemia. This approval was based on trial showing superior confirmed complete cytogenetic response by 12 months for dasatinib versus imatinib (77% versus 66%, p=0·007)[6]. However, an assessment of the drug by the National Institute of Health and Care Excellence (NICE) concluded that the estimated incremental gain in survival (22·7 years versus 21·3 years) extrapolated from the observed improvement on the surrogate endpoint came at a patient cost in excess of €200,000 per quality adjusted life years. As a result of this assessment, NICE did not recommend coverage of the drug[6].

In order for regulatory authorities and payers to use a surrogate endpoint with confidence, a validation process for such endpoints is needed. Here, we present a

three-step framework for the validation and appropriate use of surrogate endpoints in both licensing and coverage or reimbursement decisions (FIG. 1).

Establish the level of evidence. The first step is to consider the hierarchy of available evidence[7]. The biological plausibility of the relationship between the surrogate endpoint and final outcome is necessary but not sufficient. Evidence is considered to be 'level 2' when a strong correlation exists between the surrogate and the final endpoint across cohorts or at the level of the individual patient. However, individual patient correlations do not provide the highest level of evidence in order to validate surrogate measures, although they may identify good prognostic markers[8]. 'Level 1' evidence requires demonstration of the relationship between the treatment effect on the surrogate endpoint and the final outcomes, preferably across multiple randomised trials. Trial-based evidence of a final outcome is usually not available for a new healthcare technology for which surrogates are used, so this evidence needs to be sourced from other trials of the same or a similar technology — for example, trials should be of drugs from the same class or, if such evidence is not is available, drugs from a different class.

Assess the strength of the association. The second step is to assess the strength of the association between the surrogate endpoint and the final outcome. Among several approaches to address this issue, regression-based and meta-analytic approaches dominate the field. The most reliable approach is to perform a meta-analysis using patient-level data from all randomised trials of this treatment[9]. When patient-level data are available, two levels of association can be estimated: the association between the surrogate and the final outcome, and the association between the effect of treatment on the surrogate and the final outcome. Thresholds set to identify good surrogates can be as high as 0·8 for correlation coefficients (ρ) or 0·65 for coefficients of determination (R^2), which are particularly strict rules for the acceptability of putative surrogate endpoints when applied in practice[10].

Quantify the relationship between the surrogate and the final outcome. The final step relates to predicting and quantifying the effect on the final outcome based on the

observed effect on the surrogate. A quantitative approach has been proposed that consists of estimating the 'Surrogate Threshold Effect' (STE), which is the magnitude of treatment effect on the surrogate that would predict a significant treatment effect on the final outcome[11]. This is crucial for decisions on coverage and reimbursement. Regulators usually focus on early evidence of safety and efficacy to determine if the balance of benefits and risk is positive when informing the design of registration trials, while reimbursement agencies usually consider long-term effectiveness or cost-effectiveness. Whether decisions on market access and reimbursement are based on a formal economic evaluation or on the magnitude of the clinical benefit, the effect of the treatment on the surrogate endpoint needs to be large enough to predict an improvement in the final outcome of interest before the technology can be concluded to be cost-effective.

To date, few empirical assessments have investigated the adequacy of evidence for specific surrogate endpoints or groups of surrogates, particularly in terms of reimbursement policy. One exception is oncology, for which there is a long tradition of using surrogates. With few exceptions, such as metastatic colorectal or ovarian cancer, the strength of the associations between the surrogate and final outcomes tended to be relatively low in studies so far.

In conclusion, surrogates can result in market access for technologies that turn out to offer no true health benefit—or even harm—to patients and can result in overestimation of treatment effects (and economic value), which can lead to inappropriate decisions on coverage. However, the use of appropriately validated surrogate endpoints within a consistent framework provides an important potential to speed up access to innovative technologies that offer important value for patients and healthcare systems and to improve efficiency and equity within the research and development environment.

References

1. Downing NS, Krumholz HM, Ross JS & Shah ND. Characterizing the US FDA's approach to promoting transformative innovation. *Nat Rev Drug Discov* 2015; **14**:740-41.

2. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA* 2014; **311**: 368-77.

3. Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. *JAMA* 2009; **302**: 1437-43.

4. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996; **125**: 605-13.

5. Ciani O. et al. Comparison of treatment effect sizes associated with surrogate and final patient relevant outcomes in randomised controlled trials: meta-epidemiological study. *BMJ (Clinical research ed)* 2013; **346**: f457.

6. Ciani O. et al. Complete cytogenetic response and major molecular response as surrogate outcomes for overall survival in first-line treatment of chronic myelogenous leukemia: a case study for technology appraisal on the basis of surrogate outcomes evidence. *Value Health* 2013; **16**: 1081-90.

7. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA* 1999; **282**: 771-8.

8. Buyse M, Sargent DJ, Grothey A, Matheson A, de Gramont A. Biomarkers and surrogate end points--the challenge of statistical validation. *Nature reviews Clinical oncology* 2010; **7**: 309-17.

9. Burzykowski T, Molenberghs G, Buyse M. The Evaluation of Surrogate Endpoints New York: Springer Science+Business Media, Inc.; 2005.

10. Ciani O, Davis S, Tappenden P, et al. Validation of surrogate endpoints in advanced solid tumours: systematic review of statistical methods, results, and implications for policy makers. *Int J Technol Assess Health Care* 2014; **30**: 13.

11. Burzykowski T, Buyse M. Surrogate threshold effect: an alternative measure for meta-analytic surrogate endpoint validation. *Pharm Stat* 2006; **5**: 173-86.

Competing interests statement

The authors declare competing interests: see Web version for details.

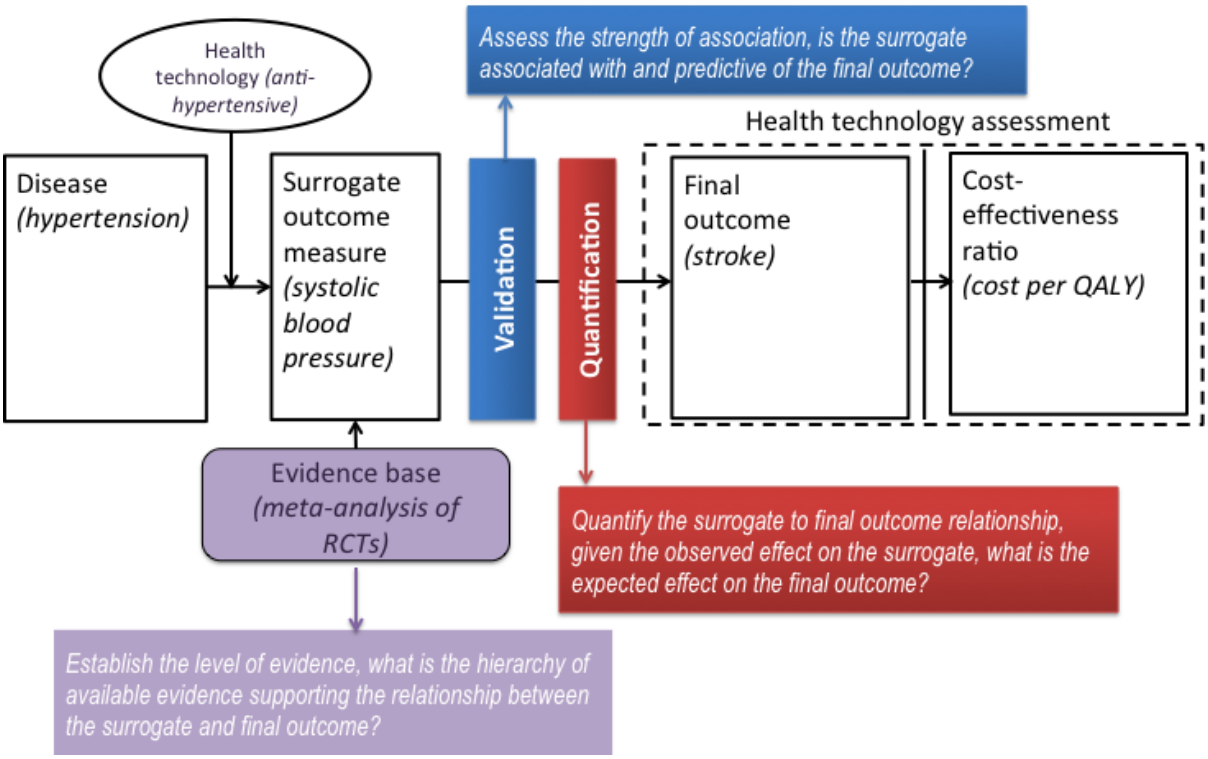
Acknowledgements

O.C. is funded by a postdoctoral research fellowship by the University of Exeter Medical School.

Disclaimer

The views expressed in this article are the personal views of the authors and may not be understood or quoted as being made on behalf of or reflecting the position of the agencies or organizations with which the authors are affiliated.

Figure 1 | **Framework for validation of surrogate endpoints: case study of anti-hypertensives. The graph shows the sequence of actions to implement in a health technology assessment of a drug technology when surrogate outcomes evidence is available. After an initial scope of the decision problem, the first step requires systematic review of the evidence explaining the relationship between the surrogate outcome and the final patient relevant outcome (*establish the level of evidence*). This evidence will then be assessed to define whether the surrogate is associated with and predictive of the final patient relevant outcome (*validation*). If so, a quantification of the estimated effect on the final outcome given the observed effect on the surrogate outcome in the setting of interest will be performed (*quantification*), and could be used as input in a cost-effectiveness analysis.**



QALY, quality-adjusted life-year; RCT, randomised controlled trial.

Online only

Competing interest statement

MB declares an association with the International Drug Development Institute Inc. and CluePoints Inc.; EDS with Dendrix Ltd.