Using surrogate endpoints in drug regulation and health technology assessment: Methodological advances and the case for change

Journal:	*BMJ*
Manuscript ID	BMJ-2021-064952
Article Type:	Analysis
BMJ Journal:	BMJ
Date Submitted by the Author:	19-Feb-2021
Complete List of Authors:	Dawoud, Dalia; National Institute for Health and Care Excellence, Science, Evidence and Analytics Directorate Naci, Huseyin; London School of Economics and Political Science, Department of Health Policy Bujkiewicz, Sylwia; University of Leicester, Department of Health Sciences Ciani, Oriana; CERGAS SDA Bocconi School of Management; University of Exeter, College of Medicine and Health
Keywords:	surrogate endpoints, pharmaceuticals, health technology assessment, drug regulation

Title:

Using surrogate endpoints in drug regulation and health technology assessment: Methodological advances and the case for change

Authors

Dalia Dawoud[1], Huseyin Naci[2*], Oriana Ciani[3,4], Sylwia Bujkiewicz[5]

[1]Scientific Adviser, Science, Evidence and Analytics Directorate, Science Policy and Research Programme, National Institute for Health and Care Excellence (NICE), London, United Kingdom

[2] Associate Professor of Health Policy, Department of Health Policy, London School of Economics and Political Science, London, United Kingdom

[3] Associate Professor of Practice, Centre for Research on Health and Social Care Management (CERGAS), SDA Bocconi, Milan, Italy

[4] College of Medicine and Health, University of Exeter, Exeter, United Kingdom

[5] Professor of Biostatistics, Biostatistics Research Group, Department of Health Sciences, University of Leicester, Leicester, United Kingdom

Corresponding Author:

Huseyin Naci, PhD
Department of Health Policy
London School of Economics and Political Science
Email: H.Naci@lse.ac.uk

ORCID ID: https://orcid.org/0000-0002-7192-5751

Standfirst

The proliferation of surrogate endpoints for regulatory approval of new drugs poses challenges for patients, clinicians, health technology assessment bodies and the wider evidence ecosystem. Dalia Dawoud and colleagues outline recent methodological advances and NICE's proposed changes to its guidance for evaluating surrogate endpoints.

Using surrogate endpoints —laboratory values, radiographic images, or other physical measures that may serve as indicators of clinical outcomes such as symptom control or mortality— in drug trials has increased substantially in recent years.[1] Surrogate endpoints have oft-cited benefits as they reduce the duration, cost, and complexity of clinical trials prior to regulatory assessment of new drugs. However, their use in clinical trials may also have unintended consequences, especially when the only evidence available at the time of new drug approval is based on surrogate endpoints.

According to the Biomarkers Definitions Working Group, a surrogate endpoint is a good predictor of clinical benefit.[2] Therefore, a putative surrogate endpoint should first be assessed for its predictive value of the treatment effect on the final clinical outcome before use in drug trials. Yet, despite important methodological advances for assessing the validity of surrogate endpoints, regulatory agencies and many health technology assessment (HTA) bodies do not routinely acknowledge the uncertainty associated with decision-making based on surrogate endpoints.

In this Analysis article, we first outline the benefits and harms of using surrogate endpoints in drug trials and describe their use for decision-making by regulatory agencies and HTA bodies. We then review the latest methodological advances for evaluating surrogate endpoints. Finally, we summarise the recently proposed changes to the NICE technology evaluation methods for using surrogate endpoints.

Benefits and harms of using surrogate endpoints in drug trials

Using surrogate endpoints can shorten the duration of clinical development for new drugs and facilitate faster patient access to innovative new therapies.[3] They, therefore, play an important role in drug research and development, at both the trial design and evaluation stage. Surrogate endpoints are particularly useful when they can provide early or more accurate measurement of a drug's effect, especially in settings where long follow-up is required before the final clinical outcome can be accurately assessed.[4] For example, this is often the case in early-stage cancer where clinical outcomes like overall survival (OS) are of primary interest to patients whilst surrogate endpoints such as event-free survival potentially can be used to measure the effect of a drug earlier.

While using surrogate endpoints in clinical trials speeds up the regulatory approval process, such endpoints can also have adverse implications for patients, clinicians, HTA bodies and the wider evidence ecosystem in health care. From the perspectives of patients and clinicians, surrogate endpoints can complicate treatment decisions;[5] they are not inherently meaningful on their own and drugs approved on the basis of surrogate endpoints may not ultimately influence patient-relevant clinical outcomes such as health-related quality of life (HRQoL) or survival. In cancer, for example, most validation studies found low correlations between surrogate endpoints and survival or HRQoL.[6–8]

Using surrogate endpoints for approving new drugs can have knock-on effects on HTA bodies, such as the National Institute for Health and Care Excellence (NICE). The assessments conducted by HTA bodies typically include comparative clinical and cost-effectiveness considerations. When new drugs receive regulatory approval based on surrogate endpoints alone, long-term effects of new drugs are associated with substantial uncertainty.

Such uncertainty may not be resolved after market entry. New trials rarely emerge in the post-marketing period to confirm a drug's clinical benefit on patient relevant outcomes such as survival and HRQoL.[9,10] Even when new trials emerge, they may fail to confirm clinical benefit. Indeed, there is a long list of drugs that were originally approved on the basis of surrogate endpoints and for which post-marketing studies failed to show evidence of clinical benefit (bevacizumab for metastatic breast cancer;[11] olaratumab for soft-tissue sarcoma;[12] hydroxyprogesterone caproate for preterm delivery,[13] among others). In some cases, drugs initially approved on the basis of surrogate endpoints were later found to be harmful (venetoclax in multiple myeloma).[14]

Use of surrogate endpoints by regulatory agencies

Pivotal studies that form the basis of regulatory approvals for new drugs frequently rely on surrogate endpoints. Over the past 3 decades, the proportion of pivotal clinical studies that collected data on surrogate endpoints has increased substantially, rising from fewer than one half in the mid-90s to approximately 60% in 2015-2017.[1] In some therapeutic areas such as cardiovascular disease and cancer, surrogate endpoints account for almost 80% of all pivotal studies that inform regulatory decisions.[15]

The recent proliferation of surrogate endpoints for drug approvals is partly due to the increase in the use of so-called 'expedited' regulatory programs that are aimed at speeding up the development, review, and approval of novel therapeutics.[16] Over the past quarter century, several expedited programs have emerged in Europe and the United States.[17] A key feature of these programs is their reliance on clinical trials with surrogate endpoints. Examples in the US include the Food and Drug Administration's (FDA) fast-track designation, breakthrough therapy designation, and priority review. Programs in Europe include the European Medicines Agency's (EMA) accelerated

assessment and Priority Medicines schemes.[18] The use of these programs has increased considerably in recent years.[19] In 2018, over three-quarters of new drugs approved by the FDA benefited from at least one expedited program. Although not all expedited programs explicitly facilitate the use of surrogate endpoints, reviews of recent approvals have shown that the use of surrogate endpoints is common across all expedited programs. For example, over 90% of cancer drugs that received the FDA's breakthrough therapy designation were approved on the basis of pivotal studies that focused on surrogate endpoints alone.[20]

The use of surrogate endpoints in certain expedited regulatory pathways is linked to "conditional" approvals where drug manufacturers are legally mandated to conduct additional trials to demonstrate the clinical benefit of their products. Even in these cases, clinical efficacy of drugs initially approved on the basis of surrogate endpoints is often subsequently "confirmed" on the basis of other surrogate endpoints.[21]

Unfortunately, regulatory guidance on the appropriate use of surrogate endpoints remains limited. In 2018, FDA published a table listing all surrogate endpoints that it has used in its assessments. However, it did not disclose whether the surrogate endpoints used in these settings have been validated. In a recent evaluation, researchers found only weak correlations between surrogate endpoints and OS in breast cancer.[22] In another systematic evaluation, researchers found that none of the surrogate endpoints used in EMA expedited approvals were validated.[23]

Use of surrogate endpoints by HTA bodies

HTA guidance on the use of surrogate endpoints is highly variable. In a recent survey of methodological guidance by 73 HTA bodies, only 40% of organisations gave specific consideration to surrogate endpoints.[24] Guidelines issued by the European network for Health Technology

Assessment (EUnetHTA)—an umbrella organisation of European HTA bodies—emphasise the need for strong evidence that an effect on the surrogate is predictive of an effect on the final outcome, preferably derived from several RCTs, in order to establish the validity of the surrogate endpoints. However, other HTA bodies such as the German Institute for Quality and Efficiency in Health Care (IQWiG); the Australian Pharmaceutical Benefits Advisory Committee (PBAC); the Portuguese National Authority of Medicines and Health Products (INFARMED) and the Canadian Agency for Drugs and Technologies in Health (CADTH), have developed more detailed prescriptive criteria for the establishing the validity of surrogate endpoints. For example, IQWiG guidance states a preference for "correlation-based" meta-analytic methods and sets a threshold for the lower bound of the confidence interval on the correlation coefficient ($R \geq 0.85$) to conclude a high correlation exists between the surrogate and final endpoint of interest.[25]

Although there has been an increase in the consideration given to the use of surrogate endpoints over the last decade, many HTA bodies still lack detailed methodological guidance for the evaluation of surrogate endpoints.[26] Among those providing methodological recommendations, there is variation in the surrogate evaluation frameworks adopted leading to potentially heterogenous conclusions about the validity of the same putative surrogate endpoints across different settings.[27]

A recent study across eight international HTA bodies, including NICE, investigated how surrogate endpoints are empirically used in clinical and cost-effectiveness analyses.[28] The level of depth and scrutiny applied by different agencies varied, with a relatively infrequent application of sound methods for surrogate validation. Overall, NICE demonstrated greater transparency than other HTA bodies when examining the strength of the evidence for the validity of putative surrogate endpoints.[29]

Validating surrogate endpoints

There is a long history of methodological efforts for evaluating the validity of surrogate endpoints. Taylor and Elston[30] proposed a three-step framework for evaluating the evidence supporting the validity of a surrogate endpoint, based on (i) biological plausibility alone, (ii) evidence of an association between the surrogate and the final endpoint at the individual patient level and (iii) evidence showing that drugs improving the treatment effect on the surrogate also improve treatment effect on the final clinical outcome across many randomised controlled trials. This framework was further extended to include a quantification of the expected effect on the final endpoint, given the observed effect on the surrogate endpoint.[31]

Evidence on patient level association is insufficient when evaluating surrogate endpoints, in particular when individual level association is based on data from a single trial,[32] as the validity confirmed based on data for one drug may not hold for others. A meta-analytic approach, based on data from a number of trials, is thus more appropriate.

Putative surrogate endpoints are validated, using meta-analytic techniques, by estimating the pattern of association between the treatment effects on surrogate and final endpoints across trials.[33–37] The recommended approach is the use of bivariate meta-analytic methods to evaluate the surrogate-to-final relationships because these methods by nature take into account not only the correlation between the treatment effects measured on these outcomes, but also all related uncertainty. The correlation, or a related measure of association, between the correlated effects is needed to quantify the surrogate relationship. Alternative criteria for validating surrogate-to-final relationship, based on the value of such measures have been proposed. These, however, are considered less informative compared to assessing the quality of predictions and associated uncertainty (see **Box 1**).[38,39]

When surrogate relationships depend on the mechanism of action of drugs or drug classes, these may be investigated in subgroups, based on data from a certain class of drugs. This may dramatically reduce the amount of available evidence for evaluation. To overcome this limitation, new methods have recently been developed. Bujkiewicz et al,[38] proposed a bivariate network meta-analytic method for surrogate endpoint evaluation which allows for modelling surrogate relationships in each treatment contrast individually whilst borrowing information from other treatment contrasts by taking into account the network structure of the data. Another method recently developed by Papanikos et al[40] extends the method by Daniels and Hughes[36] to a hierarchical model which allows for borrowing information about surrogate relationships across treatment classes. These methods have been outlines in a recent technical support document developed by NICE Decision Support Unit (DSU).[39]

Proposed changes to NICE methods of health technology evaluation in relation to surrogates

The current NICE technology appraisal methods guide, which was published in 2013, specifies that when surrogate endpoints are used to demonstrate the value of a health technology, evidence must be submitted to support the relationship between the surrogate endpoint and the final clinical outcomes, namely HRQL and survival.[41] It also stresses the importance of exploring and quantifying this relationship and the uncertainty around it. Given the methodological advances in the area of evaluating surrogate endpoints outlined above and the expanded use of these endpoints in NICE submissions, this area has been included in the Institute's currently ongoing review of its methods of health technology evaluation, initiated in 2019.[42]

The work completed as part of this review included a targeted review of the key literature on the topic of defining, evaluating and using surrogate endpoints, a review of the methods used by other HTA bodies and a Decision Support Unit report of the latest methodological advances in surrogate

endpoints' evaluation and use.[43,44] These reviews concluded that there is a case for a number of changes to be made including making the requirements for surrogate endpoint evaluation more explicit, describing the levels of evidence, the expectations for evidence of validation, and how to properly account for the uncertainty in the surrogate relationship. It was also acknowledged that a degree of flexibility would be required, while seeking the highest quality evidence possible, to accommodate a range of scenarios such as when assessing innovative or first-in-class treatments and those for rare or ultra-rare diseases.[43]

These reviews informed a number of proposed changes to NICE technology evaluation methods that aim to ensure the benefits of using surrogate endpoints, as discussed in this article, continue to be realised while avoiding the harms associated with their use. The proposals focused on the need to update that evidence requirements for surrogate endpoints to include information on the different levels of evidence, expectations for evidence of validation, and how to account for uncertainty, and should include flexibility for different evidence scenarios.[45] These proposals were included in the consultation on the changes to NICE technology evaluation methods that concluded in December 2020.[46] The outcome of this consultation will inform the next stages of NICE methods review which is due to conclude in September 2021 with the publication of a new technology evaluation program manual.[47] The main aim is to ensure that NICE evaluations continue to be robust, efficient, and based on state-of-the-art methods.[48] This is very important for supporting the NHS in securing timely access to clinically and cost-effective health technologies while ensuring that the best available evidence is used when assessing these technologies.

Conclusion

Surrogate endpoints are widely used in clinical trials with the ultimate aim of expediting research and development and achieving earlier patient access to effective treatments. This benefit is unlikely to

be achieved unless regulatory and HTA agencies are able to appropriately evaluate the proposed surrogate endpoints. Advanced analytical methods exist for this purpose and should be adopted in practice to ensure that decisions made based on surrogate endpoints achieve their intended benefits.

Key messages

- Surrogate endpoints are widely used in clinical trials with the aim of reducing the cost and time of completing these trials, but their use poses methodological challenges particularly to regulatory agencies and health technology assessment (HTA) bodies such as the National Institute for Health and Care Excellence (NICE).
- Methods for the evaluation of surrogate endpoints, including quantifying the surrogate relationship and the uncertainty around it, have developed over the years with advanced methods based on evidence synthesis approaches now available.
- These advances should be reflected in regulatory and HTA methods guidance for the evaluation of new drugs and other health technologies to ensure that decisions are made based on the best available evidence.

Acknowledgments

OC received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement #779306 (COMED—Pushing the Boundaries of Cost and Outcome Analysis of Medical Technologies). The results reflect only the author's views, and the EU is not responsible for any use that may be made of the information it contains. SB was supported by the Medical Research Council through the Methodology Research Programme grants MR/L009854/1 and MR/T025166/1. HN's work on pharmaceutical policy is supported in part by the Health Foundation.

Patient and Public Involvement

Patients or the public were not involved in the creation of our article.

Footnotes

Contributors and sources: HN devised the idea for this article. OC drafted sections relating to the HTA perspective, HN drafted sections relating to the regulatory perspective, SB drafted sections relating to the methodological and analytical perspective, DD drafted sections relating to NICE methods and proposed changes. All authors contributed to developing the first draft and writing of subsequent versions. DD is the guarantor.

Competing interests: DD is an employee of NICE. SB has served as a paid consultant providing methodological advice to NICE and Roche.

References

1 Zhang AD, Puthumana J, Downing NS, *et al.* Assessment of Clinical Trials Supporting US Food and Drug Administration Approval of Novel Therapeutic Agents, 1995-2017. *JAMA Netw Open* 2020;**3**:e203284. doi:10.1001/jamanetworkopen.2020.3284

2 Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 2001;**69**:89–95. doi:10.1067/mcp.2001.113989

3 Chen EY, Joshi SK, Tran A, *et al.* Estimation of Study Time Reduction Using Surrogate End Points Rather Than Overall Survival in Oncology Clinical Trials. *JAMA Intern Med* 2019;**179**:642–7. doi:10.1001/jamainternmed.2018.8351

4 Burzykowski T, Molenberghs G, Buyse M. *The Evaluation of Surrogate Endpoints*. Springer, New York, NY https://link.springer.com/book/10.1007/b138566#about

5 Yudkin JS, Lipska KJ, Montori VM. The idolatry of the surrogate. *BMJ* 2011;**343**:d7995. doi:10.1136/bmj.d7995

6 Prasad V, Kim C, Burotto M, *et al.* The Strength of Association Between Surrogate End Points and Survival in Oncology: A Systematic Review of Trial-Level Meta-analyses. *JAMA Internal Medicine* 2015;**175**:1389–98. doi:10.1001/jamainternmed.2015.2829

7 Hwang TJ, Gyawali B. Association between progression-free survival and patients' quality of life in cancer clinical trials. *International Journal of Cancer* 2019;**144**:1746–51. doi:10.1002/ijc.31957

8 Kovic B, Jin X, Kennedy SA, *et al.* Evaluating Progression-Free Survival as a Surrogate Outcome for Health-Related Quality of Life in Oncology: A Systematic Review and Quantitative Analysis. *JAMA Internal Medicine* 2018;**178**:1586–96. doi:10.1001/jamainternmed.2018.4710

9 Davis C, Naci H, Gurpinar E, *et al.* Availability of evidence of benefits on overall survival and quality of life of cancer drugs approved by European Medicines Agency: retrospective cohort study of drug approvals 2009-13. *BMJ* 2017;**359**:j4530. doi:10.1136/bmj.j4530

10 Pease AM, Krumholz HM, Downing NS, *et al.* Postapproval studies of drugs initially approved by the FDA on the basis of limited evidence: systematic review. *BMJ* 2017;**357**:j1680. doi:10.1136/bmj.j1680

11 Miller K, Wang M, Gralow J, *et al.* Paclitaxel plus Bevacizumab versus Paclitaxel Alone for Metastatic Breast Cancer. *N Engl J Med* 2007;**357**:2666–76. doi:10.1056/NEJMoa072113

12 Tap WD, Wagner AJ, Schöffski P, *et al.* Effect of Doxorubicin Plus Olaratumab vs Doxorubicin Plus Placebo on Survival in Patients With Advanced Soft Tissue Sarcomas: The ANNOUNCE Randomized Clinical Trial. *JAMA* 2020;**323**:1266–76. doi:10.1001/jama.2020.1707

13 Chang CY, Nguyen CP, Wesley B, *et al.* Withdrawing Approval of Makena — A Proposal from the FDA Center for Drug Evaluation and Research. *N Engl J Med* 2020;**383**:e131. doi:10.1056/NEJMp2031055

14 Kumar S, Rajkumar SV. Surrogate endpoints in randomised controlled trials: a reality check. *The Lancet* 2019;**394**:281–3. doi:10.1016/S0140-6736(19)31711-8

15 Downing NS, Aminawung JA, Shah ND, *et al.* Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005-2012. *JAMA* 2014;**311**:368–77. doi:10.1001/jama.2013.282034

16 Wallach JD, Ross JS, Naci H. The US Food and Drug Administration's expedited approval programs: Evidentiary standards, regulatory trade-offs, and potential improvements. *Clin Trials* 2018;**15**:219–29. doi:10.1177/1740774518770648

17 Darrow JJ, Avorn J, Kesselheim AS. FDA Approval and Regulation of Pharmaceuticals, 1983-2018. *JAMA* 2020;**323**:164–76. doi:10.1001/jama.2019.20288

18 Neez E, Hwang TJ, Sahoo SA, *et al.* European Medicines Agency's Priority Medicines Scheme at 2 Years: An Evaluation of Clinical Studies Supporting Eligible Drugs. *Clinical Pharmacology & Therapeutics* 2020;**107**:541–52. doi:10.1002/cpt.1669

19 Naci H, Salcher-Konrad M, Kesselheim AS, *et al.* Generating comparative evidence on new drugs and devices before approval. *The Lancet* 2020;**395**:986–97. doi:10.1016/S0140-6736(19)33178-2

20 Puthumana J, Wallach JD, Ross JS. Clinical Trial Evidence Supporting FDA Approval of Drugs Granted Breakthrough Therapy Designation. *JAMA* 2018;**320**:301–3. doi:10.1001/jama.2018.7619

21 Naci H, Smalley KR, Kesselheim AS. Characteristics of Preapproval and Postapproval Studies for Drugs Granted Accelerated Approval by the US Food and Drug Administration. *JAMA* 2017;**318**:626–36. doi:10.1001/jama.2017.9415

22 Gyawali B, Hey SP, Kesselheim AS. Evaluating the evidence behind the surrogate measures included in the FDA's table of surrogate endpoints as supporting approval of cancer drugs. *EClinicalMedicine* 2020;**21**. doi:10.1016/j.eclinm.2020.100332

23 Schuster Bruce C, Brhlikova P, Heath J, *et al.* The use of validated and nonvalidated surrogate endpoints in two European Medicines Agency expedited approval pathways: A cross-sectional study of products authorised 2011–2018. *PLOS Medicine* 2019;**16**:e1002873. doi:10.1371/journal.pmed.1002873

24 Grigore B, Ciani O, Dams F, *et al.* Surrogate Endpoints in Health Technology Assessment: An International Review of Methodological Guidelines. *PharmacoEconomics* 2020;**38**:1055–70. doi:10.1007/s40273-020-00935-1

25 Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen, (IQWiG). Aussagekraft von surrogatendpunkten in der onkologie., Institut fuer Qualitaet und Wirtschaftlichkeit im Gesundheitswesen (IQWiG). Validity of surrogate parameters in oncology (Rapid report). Cologne: 2011.

26 Garrido MV, Mangiapane S. Surrogate outcomes in health technology assessment: An international comparison. *International Journal of Technology Assessment in Health Care* 2009;**25**:315–22. doi:10.1017/S0266462309990213

27 Ciani O, Davis S, Tappenden P, *et al.* Validation of surrogate endpoints in advanced solid tumors: systematic review of statistical methods, results, and implications for policy makers. *Int J Technol Assess Health Care* 2014;**30**:312–24. doi:10.1017/S0266462314000300

28 Ciani O, Grigore B, Blommestein H, *et al.* Validity of surrogate endpoints and their impact on coverage recommendations. A retrospective analysis across international health technology assessment agencies. *Med Decis Making* 2021.

29 Oriana C, Martin H, Toby P, *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 in Health* 2013;**16**:1081–90. doi:10.1016/j.jval.2013.07.004

30 Taylor R, Elston J. The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports. *Health Technol Assess* 2009;**13**:8. doi:10.3310/hta13080

31 Ciani O, Buyse M, Drummond M, *et al.* Use of surrogate end points in healthcare policy: a proposal for adoption of a validation framework. *Nature Reviews Drug Discovery* 2016;**15**:516–516. doi:10.1038/nrd.2016.81

32 Fleming T, DeMets D. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;**125**:605–13.

33 Bujkiewicz S, Thompson JR, Spata E, *et al.* Uncertainty in the Bayesian meta-analysis of normally distributed surrogate endpoints. *Stat Methods Med Res* 2015;**26**:2287–318. doi:10.1177/0962280215597260

34 Bujkiewicz S, Thompson JR, Riley RD, *et al.* Bayesian meta-analytical methods to incorporate multiple surrogate endpoints in drug development process. *Statistics in Medicine* 2016;**35**:1063–89. doi:10.1002/sim.6776

35 Burzykowski T, Molenberghs G, Buyse M, *et al.* Validation of surrogate end points in multiple randomized clinical trials with failure time end points. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 2001;**50**:405–22. doi:10.1111/1467-9876.00244

36 Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. *Statistics in Medicine* 1997;**16**:1965–82. doi:10.1002/(SICI)1097-0258(19970915)16:17<1965::AID-SIM630>3.0.CO;2-M

37 Buyse M, Molenberghs G, Burzykowski T, *et al.* The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics* 2000;**1**:49–67. doi:10.1093/biostatistics/1.1.49

38 Bujkiewicz S, Jackson D, Thompson JR, *et al.* Bivariate network meta-analysis for surrogate endpoint evaluation. *Statistics in Medicine* 2019;**38**:3322–41. doi:10.1002/sim.8187

39 Bujkiewicz S, Achana F, Papanikos T, *et al.* NICE DSU Technical Support Document 20: Multivariate meta-analysis of summary data for combining treatment effects on correlated outcomes and evaluating surrogate endpoints. Sheffield: : ScHARR, University of Sheffield 2019. http://nicedsu.org.uk/wp-content/uploads/2020/10/TSD-20-mvmeta-final.pdf

40 Papanikos T, Thompson JR, Abrams KR, *et al.* Bayesian hierarchical meta-analytic methods for modeling surrogate relationships that vary across treatment classes using aggregate data. *Statistics in Medicine* 2020;**39**:1103–24. doi:10.1002/sim.8465

41 National Institute for Health and Care Excellence (NICE). Guide to the methods of technology appraisal 2013. 2013.https://www.nice.org.uk/process/pmg9/chapter/foreword (accessed 22 Dec 2020).

42 National Institute for Health and Care Excellence (NICE). Centre for Health Technology Evaluation (CHTE) methods review – task and finish group specifications. 2019. https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation (accessed 22 Dec 2020).

43 National Institute for Health and Crae Excellence (NICE). CHTE methods review: Sources and synthesis of evidence Task and finish group report. 2020. https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation

44 Welton N, Phillippo D, Owen R, *et al.* CHTE2020 Sources and Synthesis of Evidence: Update to Evidence Synthesis Methods Report by The Decision Support Unit. Sheffield: : ScHARR, University of Sheffield 2020. http://nicedsu.org.uk/wp-content/uploads/2020/11/CHTE-2020_final_20April2020_final.pdf

45 National Institute for Health and Crae Excellence (NICE). The NICE methods of health technology evaluation: The case for change. 2020. https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation

46 National Institute for Health and Crae Excellence (NICE). Reviewing our methods for health technology evaluation: Consultation. 2020.https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/chte-methods-consultation

47 National Institute for Health and Crae Excellence (NICE). Changes we're making to health technology evaluation. 2019.https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/changes-to-health-technology-evaluation (accessed 13 Jan 2021).

48 Boysen M, Watson I. Reflections on NICE's Uptake of New Methods: Past, Present, and the 2020 Review. *PharmacoEconomics* 2020;**38**:243–5. doi:10.1007/s40273-020-00888-5

Box 1. Alternative criteria for surrogate endpoint evaluation:

- A number of approaches and criteria have been proposed for assessing the validity of a surrogate endpoint. For study-level surrogacy, these include a correlation between treatment effects on the surrogate endpoint and the final outcome, regression parameters describing the association between these effects (intercept, slope and conditional variance) or study-level R-squared.

- Authors and policy makers have made a number of recommendations regarding the values of these parameters, but also have acknowledged that a decision on the use of a surrogate endpoint in policy decision making may result from a balance between the strength of evidence of the surrogate relationship and the unmet need, for example when new treatment under consideration is in a priority disease area.

- More recently, Bujkiewicz et al[36] and the NICE Decision Support Unit (DSU) report[37] noted that it is difficult to quantify how large the correlation should be in order to consider the surrogate endpoint suitable to make the prediction. Instead, the strength (or weakness) of the surrogate relationship will manifest itself in the uncertainty around the predicted treatment effect on the final outcome. A weaker surrogate relationship will result in a larger interval and hence increased uncertainty about the regulatory or clinical decision made based on such prediction.