Raising the bar for using surrogate endpoints in drug regulation and health technology assessment

Journal:	*BMJ*
Manuscript ID	BMJ-2021-064952.R1
Article Type:	Analysis
Date Submitted by the Author:	07-Jul-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 Ciani, Oriana; CERGAS SDA Bocconi School of Management; University of Exeter, College of Medicine and Health Bujkiewicz, Sylwia; University of Leicester, Department of Health Sciences
Keywords:	surrogate endpoints, pharmaceuticals, health technology assessment, drug regulation

Title:

Raising the bar for using surrogate endpoints in drug regulation and health technology assessment

Authors:

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

[1] Senior 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

Associate Professor of Health Policy

Department of Health Policy

London School of Economics and Political Science

Houghton Street

London, WC2A 2AE

United Kingdom

Phone: +44 (0) 20 7955 6874

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 argue for raising the evidence standards for using surrogate endpoints by regulatory agencies and health technology assessment bodies.

On 7 June 2021, the US Food and Drug Administration (FDA) granted accelerated approval to aducanumab for the treatment of Alzheimer's Disease. The FDA based its decision on the drug's amyloid-reducing effects despite evidence from several earlier studies that shrinkage of beta-amyloid protein plaques does not predictably delay cognitive impairment in patients. [1]The decision has drawn significant attention to the use of 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 clinical trials of new drugs.[2] In fact, the approval of aducanumab is only the latest example of growing regulatory reliance on surrogate endpoints. Using surrogate endpoints in drug trials has increased substantially in recent years.[3]

According to the Biomarkers Definitions Working Group, a surrogate endpoint is "a good predictor of clinical benefit."[4] Therefore, a candidate 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 rarely evaluate surrogate endpoints, and do not routinely acknowledge the uncertainty associated with their use in decision-making.

In this article, we first review the benefits and harms of using surrogate endpoints in drug trials and describe their use for decision-making by regulatory agencies and HTA bodies. After outlining the key methodological considerations for evaluating surrogate endpoints, we discuss how regulators and HTA bodies can raise the bar for using surrogate endpoints.

Benefits and harms of using surrogate endpoints in drug trials

Using surrogate endpoints can reduce the duration, cost, and complexity of clinical trials prior to regulatory assessment, and facilitate faster patient access to new therapies.[5] Surrogate endpoints are useful in chronic disease settings when they can provide early and accurate measurement of a drug's effect, especially when long follow-up is required before the final clinical outcome can be

assessed.[6] For example, in early-stage gastric cancer, clinical outcomes like overall survival—how long patients live after receiving treatment—are of primary interest to patients whilst surrogate endpoints such as disease-free survival potentially can be used to measure drug effects earlier.[7] By contrast, surrogate endpoints are not useful when a drug's effect on the final clinical outcome can be observed within a relatively short time frame, e.g., in acute conditions.[8]

While using surrogate endpoints in clinical trials can speed up the regulatory approval process, such endpoints can also have adverse implications for patients, clinicians, HTA bodies and the wider evidence ecosystem. From the perspectives of patients and clinicians, surrogate endpoints can complicate treatment decisions;[9] 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.[10–12]

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.[13–15] 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.[16] Examples include bevacizumab for metastatic breast cancer;[17] olaratumab for soft-tissue sarcoma;[18] hydroxyprogesterone caproate for preterm delivery,[19] among others. In some cases, drugs initially approved on the basis of surrogate endpoints were later found to be harmful (venetoclax in multiple myeloma).[20]

Use of surrogate endpoints by regulators

Pivotal studies that form the basis of regulatory approvals for new drugs frequently use surrogate endpoints alone. Over the past 3 decades, the proportion of pivotal clinical studies that collected data

on surrogate endpoints has increased, rising from fewer than one half in the mid-90s to approximately 60% in 2015-2017.[3] In some therapeutic areas such as cancer, surrogate endpoints account for almost 80% of all pivotal studies.[21]

The recent proliferation of surrogate endpoints is partly due to the increase in the use of 'expedited' regulatory programs that are aimed at speeding up the development, review, and approval of drugs.[22] Over the past quarter century, lobbying by pharmaceutical companies has put pressure on policymakers to establish several expedited programs in Europe and the United States.[23] These programs also meet perceived patient demand for faster access to potentially effective therapies. Examples in the US include the FDA accelerated approval pathway and breakthrough therapy designation. Programs in Europe include the European Medicines Agency's (EMA) accelerated assessment and Priority Medicines schemes.[24] The use of these programs has increased considerably.[25] In 2018, over three-quarters of new drugs approved by the FDA benefited from at least one expedited program.[23]

Not all expedited programs explicitly facilitate the use of surrogate endpoints. However, 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 measured surrogate endpoints alone.[26]

The use of surrogate endpoints in certain expedited regulatory programs like the FDA's accelerated approval pathway is linked to "conditional" approvals where drug manufacturers are legally mandated to conduct additional trials to demonstrate the clinical benefit of their products. Even when post-approval studies are required, clinical efficacy of drugs initially approved on the basis of surrogate endpoints is often subsequently "confirmed" on the basis of other surrogate endpoints.[27,28]

Unfortunately, surrogate endpoints accepted by regulators are not routinely validated. In a recent study, researchers found only weak correlations between surrogate endpoints and survival in breast cancer.[29] In another analysis, researchers found that none of the surrogate endpoints used in EMA expedited approvals were validated.[30]

Use of surrogate endpoints by HTA bodies

HTA guidance on the use of surrogate endpoints is highly variable [31]. In a recent survey of methodological guidance by 73 HTA bodies, only 40% gave specific consideration to surrogate endpoints.[32] Guidelines issued by the European network for Health Technology Assessment—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 clinical outcome, preferably derived from several randomized trials.

However, some HTA bodies such as the German Institute for Quality and Efficiency in Health Care (IQWiG) and the Australian Pharmaceutical Benefits Advisory Committee have developed more prescriptive criteria for establishing the validity of surrogate endpoints. For example, IQWiG prefers "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 clinical outcome.[33]

A recent study investigated how surrogate endpoints are used in clinical and cost-effectiveness analyses by eight HTA bodies.[34] The level of depth and scrutiny applied by different organisations varied, with an infrequent consideration of the level of available evidence and application of statistical methods for surrogate validation.[35] Such variation across HTA bodies yields potentially heterogenous conclusions about the validity of the same putative surrogate endpoints across different settings.[36]

Validating surrogate endpoints

There is a long history of methodological efforts for evaluating the validity of surrogate endpoints. Taylor and Elston [37] proposed a three-step framework for evaluating surrogate endpoints, based on (i) biological plausibility alone, (ii) evidence of an association between the surrogate and the clinical 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 trials. This framework was further extended to quantify the expected effect of the surrogate on the final clinical outcome.[38]

Evidence at the individual patient level alone is insufficient to evaluate surrogate endpoints especially when such evidence is obtained from a single trial.[39] This is because the observed surrogate-to-clinical outcome relationship for one drug may not hold for another. Fleming and DeMets describe a number of scenarios where, depending on the mechanism of action of a treatment,

a surrogate may not fully mediate the treatment effect on the clinical outcome, as different interventions may affect disease pathways in different ways, involving sometimes one endpoint but not the other. [39] Meta-analysis, which combines data from a number of randomised trials, is therefore more appropriate for evaluating the association between the treatment effects on the candidate surrogate endpoint and on the final clinical outcome.[40]

There is growing methodological consensus for using bivariate meta-analysis methods to evaluate the surrogate-to-final relationships. [41–45] These methods take into account not only the correlation between the treatment effects, but also uncertainty. The correlation between treatment effects on the surrogate endpoint and the final outcome, or a related measure of association between the correlated effects is needed to quantify the surrogate relationship.[44,45] Table 1 lists examples of candidate surrogate endpoints evaluated using meta-analysis methods.

When the amount of available evidence from clinical trials is limited (e.g., for drugs targeting genetic biomarkers in small patient populations), novel bivariate network meta-analysis methods, [46] or hierarchical models,[47] allow for using all available data on similar drugs or drug classes. These advanced methods are highlighted in reports prepared by the NICE Decision Support Unit.[48,49]

Way forward

Regulators' evidence standards determine the quantity and quality of evidence available on new drugs. During clinical development, manufacturers engage with regulators to agree on the designs of studies (including endpoints) that will be used to support regulatory assessment. Regulators therefore have an opportunity to adopt rigorous criteria for accepting surrogate endpoints. Using surrogate endpoints should only be considered in chronic disease settings with substantial unmet need, especially when collecting data on clinical outcomes requires clinical trials with unfeasibly long follow up durations. When generating evidence on clinical outcomes is not feasible, regulators should routinely evaluate the validity of the proposed surrogate endpoints.

Raising the bar for the regulatory use of surrogate endpoints may increase the cost and duration of drug development. In cancer, trials collecting data on clinical outcomes lasted on average 11 months longer than those measuring surrogate endpoints.[5] However, this need not hamper pharmaceutical innovation. In the past, regulatory guidance encouraging manufacturers to evaluate the cardiovascular outcomes of anti-diabetic medications incentivised the generation of patient-centred evidence without adversely affecting research and development.[50,51]

Historically, the onus to evaluate surrogate endpoints has, thus, fallen on HTA bodies. For example, NICE has recently proposed changes to its methods of health technology assessment to strengthen the evidence requirements for the use of surrogate endpoints, while still allowing flexibility when desired evidence is not available (**Box 1**).[52,53] Involving HTA bodies in early regulatory interactions with manufacturers may help align evidence requirements on surrogate endpoints. The UK Innovative Licensing and Access Pathway managed by the Medicines and Healthcare Products Regulatory Agency, NICE and the Scottish Medicines Consortium is aimed at facilitating such alignment.[54]

Ultimately, regulatory and HTA decisions regarding the use of surrogate endpoints need to weigh the strength of available evidence on the validity of surrogates alongside other considerations such as unmet therapeutic need. When making such trade-offs, quantifying how well a candidate surrogate predicts the final clinical outcome can provide valuable information.[46,48] If recommended meta-analysis methods are used for validation, the strength (or weakness) of the surrogate will be reflected in the uncertainty around the predicted treatment effect on the final outcome. A weaker surrogate will yield a larger interval and hence greater uncertainty.

Greater involvement of patients (and organisations representing patients) in regulatory and HTA processes is also essential to ensure that uncertainty related to using surrogate endpoints is explicitly presented and taken into account. Earlier studies showed that patients often overestimate the benefits of treatments on the basis of surrogate endpoints.[55] Patient input can help guide regulatory and HTA decisions regarding the appropriate use of surrogate endpoints.

Conclusion

Surrogate endpoints are widely used in clinical trials supporting regulatory and HTA assessments of new drugs. Despite their oft-cited benefits, using surrogate endpoints often complicates decision-making in clinical practice and health policy. The harms of using surrogate endpoints may outweigh the intended benefits if they are not valid predictors of clinical benefit. When generating evidence on clinical outcomes is not feasible, regulators and HTA bodies should routinely examine the performance of surrogate endpoints, acknowledge and quantify uncertainty according to an established evaluation framework.

Key messages

- Surrogate endpoints are widely used by regulators to expedite the approval of new drugs, but their use introduces substantial uncertainty about the drugs' clinical and cost-effectiveness.
- Regulators should be more selective when accepting surrogate endpoints in clinical trials as their decisions can have knock-on effects on value assessment.
- When generating data on clinical outcomes is not feasible, regulators and health technology assessment bodies should routinely evaluate the validity of surrogate endpoints according to a common evaluation framework.
- Despite the availability of methods to quantify the surrogate-to-final relationship and its associated uncertainty (and to predict the treatment effect on the final clinical outcome), they are rarely used by regulators and health technology assessment bodies.

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). Issues discussed in this paper 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 Panel grants MR/L009854/1 and MR/T025166/1.

Footnotes

Contributors and sources: DD is an expert on health technology assessment methods research and has been involved in the ongoing update of NICE's health technology evaluation methods. HN's research examines the evidence supporting regulatory decisions on drugs in the US and Europe. OC has written extensively on the role of surrogate endpoints in health care policy and cost-effectiveness models. She previously contributed to the development of surrogate validation frameworks. SB's expertise is in Bayesian evidence synthesis methods. She has developed novel methods for modelling surrogate endpoints, which are proposed to be included NICE's update of its methods guide. HN devised the idea for this article. 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. The views expressed are those of DD and not those of NICE. SB has served as a paid consultant providing methodological advice to NICE, Roche and RTI Health Solutions. SB has previously received research funding from European Federation of

Pharmaceutical Industries & Associations (EFPIA) as part of unrelated European Union IMI GetReal project. HN previously received funding from the Pharmaceutical Group of the European Union for an unrelated systematic review on community pharmacists. HN currently receives funding from the Health Foundation on an unrelated project on pharmaceutical policy.

Licence for Publication

References

1 Ackley SF, Zimmerman SC, Brenowitz WD, *et al.* Effect of reductions in amyloid levels on cognitive change in randomized trials: instrumental variable meta-analysis. *BMJ* 2021;**372**:n156. doi:10.1136/bmj.n156

2 Alexander GC, Emerson S, Kesselheim AS. Evaluation of Aducanumab for Alzheimer Disease: Scientific Evidence and Regulatory Review Involving Efficacy, Safety, and Futility. *JAMA* 2021;**325**:1717–8. doi:10.1001/jama.2021.3854

3 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

4 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

5 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

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

7 Oba K, Paoletti X, Alberts S, *et al.* Disease-free survival as a surrogate for overall survival in adjuvant trials of gastric cancer: a meta-analysis. *J Natl Cancer Inst* 2013;**105**:1600–7. doi:10.1093/jnci/djt270

8 Institute of Medicine (US) Committee on Qualification of Biomarkers and Surrogate Endpoints in Chronic Disease, Micheel C, Ball J. *Evaluation of Biomarkers and Surrogate Endpoints in Chronic Disease.* Washington (DC): : National Academies Press (US) 2010. https://www.ncbi.nlm.nih.gov/books/NBK220297/ doi: 10.17226/12869 (accessed 5 Jul 2021).

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

10 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

11 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

12 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

13 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

14 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

15 Wallach JD, Ciani O, Pease AM, *et al.* Comparison of treatment effect sizes from pivotal and postapproval trials of novel therapeutics approved by the FDA based on surrogate markers of disease: a meta-epidemiological study. *BMC Med* 2018;**16**:45. doi:10.1186/s12916-018-1023-9

16 Svensson S, Menkes DB, Lexchin J. Surrogate outcomes in clinical trials: a cautionary tale. *JAMA Intern Med* 2013;**173**:611–2. doi:10.1001/jamainternmed.2013.3037

17 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

18 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

19 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

20 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

21 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

22 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

23 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

24 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

25 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

26 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

27 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

28 Gyawali B, Hey SP, Kesselheim AS. Assessment of the Clinical Benefit of Cancer Drugs Receiving Accelerated Approval. *JAMA Intern Med* 2019;**179**:906–13. doi:10.1001/jamainternmed.2019.0462

29 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

30 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

31 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

32 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

33 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.

34 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;**(In press)**.

35 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

36 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

37 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

38 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

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

40 Joffe MM, Greene T. Related Causal Frameworks for Surrogate Outcomes. *Biometrics* 2009;**65**:530–8. doi:https://doi.org/10.1111/j.1541-0420.2008.01106.x

41 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

42 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

43 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

44 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

45 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

46 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

47 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

48 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

49 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

50 Sharma A, Pagidipati NJ, Califf RM, *et al.* Impact of Regulatory Guidance on Evaluating Cardiovascular Risk of New Glucose-Lowering Therapies to Treat Type 2 Diabetes Mellitus: Lessons Learned and Future Directions. *Circulation* 2020;**141**:843–62. doi:10.1161/CIRCULATIONAHA.119.041022

51 Hwang TJ, Franklin JM, Kesselheim AS. Effect of US Food and Drug Administration's Cardiovascular Safety Guidance on Diabetes Drug Development. *Clin Pharmacol Ther* 2017;**102**:290–6. doi:10.1002/cpt.705

52 National Institute for Health and Care 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

53 National Institute for Health and Care 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

54 Medicines and Healthcare products Regulatory Agency. Innovative Licensing and Access Pathway. 2021. https://www.gov.uk/guidance/innovative-licensing-and-access-pathway

55 Raphael MJ, Robinson A, Booth CM, *et al.* The Value of Progression-Free Survival as a Treatment End Point Among Patients With Advanced Cancer: A Systematic Review and Qualitative Assessment of the Literature. *JAMA Oncol* 2019;**5**:1779–89. doi:10.1001/jamaoncol.2019.3338

56 Buyse M, Burzykowski T, Carroll K, *et al.* Progression-Free Survival Is a Surrogate for Survival in Advanced Colorectal Cancer. *JCO* 2007;**25**:5218–24. doi:10.1200/JCO.2007.11.8836

57 Sormani MP, Bonzano L, Roccatagliata L, *et al.* Surrogate endpoints for EDSS worsening in multiple sclerosis. *Neurology* 2010;**75**:302–9. doi:10.1212/WNL.0b013e3181ea15aa

58 Inker LA, Mondal H, Greene T, *et al.* Early Change in Urine Protein as a Surrogate End Point in Studies of IgA Nephropathy: An Individual-Patient Meta-analysis. *Am J Kidney Dis* 2016;**68**:392–401. doi:10.1053/j.ajkd.2016.02.042

59 Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *The Lancet* 2005;**366**:1267–78. doi:10.1016/S0140-6736(05)67394-1

60 Lassere M, Johnson K, Boers M, *et al.* Definitions and validation criteria for biomarkers and surrogate endpoints: development and testing of a quantitative hierarchical levels of evidence schema. *J Rheumatol* 2007;**34**:607–15.

61 Ciani O, Buyse M, Drummond M, *et al.* Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. *Value in Health* 2017;**20**:487–95. doi:10.1016/j.jval.2016.10.011

Table 1: Examples of candidate surrogate endpoints evaluated using meta-analysis

Disease area	Candidate surrogate endpoint	Final clinical outcome	Reference
Advanced colorectal cancer	Progression-free survival	Overall survival	[56]
Gastric cancer	Disease-free survival	Overall survival	[7]
Multiple sclerosis	Relapse rate	Disability progression	[57]
Immunoglobulin A nephropathy	Change in proteinuria	Doubling of serum creatinine level, end-stage renal disease, or death	[58]
Cardiovascular disease	Low-density lipoprotein	Major coronary events	[59]
	Diastolic blood pressure	Stroke	[60]

Box 1. Proposed changes to NICE Methods Guide:[53]

Proposed change*	**Rationale**
Addition of: Three levels of evidence for surrogate relationships can be considered in decision making [61]: • Level 3: biological plausibility of relation between surrogate and final outcome • Level 2: consistent association between surrogate and final outcome. This would usually be derived from epidemiological or observational studies. • Level 1: Treatment effect on surrogate corresponds to commensurate treatment effect on the final outcome as shown in RCTs.	The addition of this hierarchy will clarify the different levels of evidence for a surrogate relationship. Currently, level 3 and 2 are the most commonly seen and used in submissions to NICE. Based on the literature and the Decision Support Unit (DSU) reports, these are not on their own sufficient to support the surrogacy relationship for the purpose of Health Technology Appraisal (HTA) and use in economic modelling. Level 1 evidence (RCTs) is the most appropriate for this purpose, but it is recognised that this level of evidence will not always be available.
Addition of: For a surrogate outcome to be considered validated, there needs to be good evidence that the relative effect of a technology on the surrogate is predictive of its relative effect on the final outcome. This evidence should be obtained from a meta-analysis of level 1 evidence (that is, RCTs) that reported treatment effects on both the surrogate and the final outcomes, using the recommended meta-analytic methods outlined in TSD20 (bivariate meta-analytic methods). If other levels of evidence are used to support the surrogate relationship, this will need to be justified.	With an increasing number of new surrogate endpoints being proposed and a shift towards NICE reviewing earlier data, it is recommended that the evidence requirements around the evaluation of surrogate endpoints be made more explicit in the NICE methods guide. Robust meta-analytic methods are currently available to conduct such analysis, including the bivariate meta-analytic methods proposed by the NICE DSU.
Addition of: The validation of a surrogate outcome is specific to the population and technology type under consideration.	The majority of the evidence reviewed supported this. Other HTA bodies have specified this as a requirement.
Addition of: • Extrapolation of a surrogate-to-final relationship to a different population or technology of a different class or with a different mechanism of action needs thorough justification. • It should be undertaken using the recommended meta-analytic methods that allow borrowing of information from sufficiently similar treatment classes, populations, and treatment settings, as outlined in TSD20. Existing relevant meta-analytic models may be used. However, when historical models are based on data collected in a different setting, development of a new model, using appropriate meta-analytic techniques, is recommended. This may include network meta-analysis or	If a company wants to use evidence of a surrogate to final outcome relationship from a different population or technology type, then thorough justification should be made explicitly. This may be needed for rare conditions or first-in-class treatments, where a meta-analysis of RCTs is not available to support the surrogate endpoint. Borrowing of information from other treatment classes that are likely to have similar or relevant mechanism of action is possible as methodologies exist for this, based on the NICE DSU report.

Proposed change*	Rationale
hierarchical methods reflecting differences in mechanism of action between treatment classes or for first-in-class scenarios.	
Amendment: **Current wording** "In all cases, the uncertainty associated with the relationship between the endpoint and health-related quality of life or survival should be explored and quantified." **Proposed wording:** "In all cases, the uncertainty associated with the relationship between the surrogate endpoints and the final outcomes should be quantified and presented. It should also be included through probabilistic sensitivity analysis and can be further explored in scenario analysis."	Sensitivity analyses submitted to NICE rarely include the uncertainty around the surrogate to final relationship, as they simply include the uncertainty around the estimate of treatment effect on the surrogate outcome. This important source of uncertainty should be explicitly noted in the NICE methods guide to ensure the committees are presented with a true representation of the uncertainty around the final cost-effectiveness estimate. This has also been included in the Canadian Agency for Drugs and Technologies in Health and Pharmaceutical Benefits Advisory Committee methods guides. The specification that the surrogate relationship uncertainty must be included within the model is particularly important due to the increasing number of new surrogate endpoints being proposed and a shift towards NICE reviewing earlier data.

NICE: National Institute for Health and Care Excellence; RCT: randomised controlled trial; TSD: technical support document; DSU: Decision Support Unit; NMA: network meta-analysis; HTA: health technology assessment These proposed changes were put out for public consultation in 2020. The final wording in the updated methods guide might be different to those listed above.*