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 major 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 to measure whether a new drug works can reduce the duration, cost, and complexity of clinical trials prior to regulatory assessment, and facilitate faster patient access to new therapies, especially in chronic disease settings.[3] 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.[4] In a recent evaluation, using surrogate endpoints in cancer drug trials reduced clinical development time by approximately 11 months compared with measuring overall survival. [3] However, the use of such endpoints can also have negative implications.

Regulatory reliance on surrogate endpoints makes it challenging for HTA bodies, such as the National Institute for Health and Care Excellence (NICE), to make their decisions. 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, assessing how well they work in terms of impact on patient-relevant clinical outcomes, such as health-related quality of life and survival, in the short and long term are fraught with considerable uncertainty.

For patients and clinicians, surrogate endpoints can complicate treatment decisions.[5] Surrogate endpoints are not inherently meaningful on their own, and clinicians and patients may misinterpret

drug effects on surrogate endpoints as clinically meaningful improvements.[6] This matters, because drugs approved on the basis of surrogate endpoints may not ultimately influence patient-relevant clinical outcomes. In cancer, for example, most approved drugs with effects on surrogate endpoints such as response rates and progression-free survival (that were imagined to be predictive of patient-relevant benefit) do not, in fact, improve quality of life or prolong survival.[7–9]

There is a long history of drugs that were originally approved on the basis of surrogate endpoints and for which later studies failed to show evidence of clinical benefit.[10] An oft-cited example is bevacizumab for metastatic breast cancer.[11] In 2008, FDA granted accelerated approval to bevacizumab for the treatment of metastatic breast cancer based on its early effects on a surrogate endpoint, progression-free survival. In 2011, FDA revoked its approval for bevacizumab's metastatic breast cancer indication when clinical trials failed to show that patients receiving bevacizumab lived longer than those receiving control treatment.

Other examples include olaratumab, which extended progression-free survival but did not prolong survival for patients with soft-tissue sarcoma,[12] hydroxyprogesterone caproate, which effectively reduced the risk of recurrent births but did not improve neonatal outcomes,[13] and atezolizumab, which achieved a higher response rate compared to control but did not extend overall survival in patients with urothelial carcinoma .[14] In some cases, drugs initially approved on the basis of surrogate endpoints were later found to be harmful. For example, patients with multiple myeloma who received venetoclax had shorter survival than those who received a control treatment, despite evidence suggesting that venetoclax was more effective than control on the basis of progression-free survival).[15]

In this article, we argue for more selective use of surrogate endpoints when evaluating new drugs. Surrogate endpoints should only be used in chronic disease settings, especially when collecting data on patient-relevant clinical outcomes requires trials with unattainably long follow up durations. When generating direct evidence on patient-relevant clinical outcomes is not possible, decision-makers should systematically evaluate the relationship between surrogate endpoints and clinical outcomes.

Regulatory enthusiasm for surrogate endpoints

Over the past 3 decades, the proportion of clinical studies measuring the efficacy of new drugs via surrogate endpoints alone has increased, rising from fewer than one half in the mid-90s to

approximately 60% in 2015-2017.[16] In some therapeutic areas such as cancer, surrogate endpoints account for almost 80% of all clinical studies supporting regulatory approvals.[17] This means that in some therapeutic areas, only a minority of new drugs are now approved on the basis of evidence that they improve how patients feel or function, or how long they live.

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.[18] 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.[19] These programs also meet perceived patient demand for faster access to potentially effective therapies in therapeutic areas with significant unmet needs. In the US, the FDA "accelerated approval" pathway was established at the height of the HIV/AIDS crisis in the early 1990s. Other examples of expedited programs in the US include the "breakthrough therapy," "priority review," and "fast track" designations. Programs in Europe include the European Medicines Agency's (EMA) "accelerated assessment" and "Priority Medicines" schemes.[20]

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, however, clinical efficacy of drugs initially approved on the basis of surrogate endpoints is often subsequently "confirmed" on the basis of other surrogate endpoints.[21,22] For example, both pre-approval and mandated post-approval studies supporting FDA's accelerated approval of lapatinib (for the treatment of postmenopausal women with HER2-positive metastatic breast cancer) tested surrogate endpoints.[21] This practice may meet regulators' expectations but falls far short of reliable evidence of patient benefit.

Limited guidance from regulators and HTA bodies

There is little consensus for defining a "valid" surrogate, as it is difficult to set specific thresholds to grade the strength of association with the final clinical outcome. Yet, some organisations such as the German Institute for Quality and Efficiency in Health Care (IQWiG) have prescriptive criteria for accepting surrogate endpoints. IQWiG 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.[23] Most other agencies have no similar cut-offs for accepting surrogate endpoints.

There is actually a long history of methodological efforts for evaluating surrogate endpoints. In 2009, Taylor and Elston [24] recommended a three-step framework, based on (i) biological plausibility alone, (ii) evidence of an observational association between the surrogate and the clinical endpoint at the individual patient level and (iii) evidence from multiple randomised trials showing that drugs improving the treatment effect on the surrogate also improve treatment effect on the final clinical outcome. This framework was further extended to quantify the expected treatment effect on the final clinical outcome based on the surrogate.[25]

However, this framework is rarely used by regulatory agencies. In 2018, FDA published a table listing all surrogate endpoints that it has used in its assessments without disclosing any information about their usefulness in predicting clinical benefit.[26] Academic researchers are increasingly filling this evidence gap and examining the strength of the association between surrogate endpoints that are commonly used by regulators and patient-relevant clinical outcomes. [27,28] In a recent study, researchers found only weak or missing correlations between surrogate endpoints and survival in breast cancer using the Taylor and Elston framework.[29] In another analysis, researchers found that none of the surrogate endpoints used in EMA expedited approvals were evaluated in independent studies.[30]

Similarly, HTA bodies rarely use this framework to evaluate surrogate endpoints,[31] Indeed, HTA guidance on the use of surrogate endpoints has been highly variable [32]. In a recent survey of methodological guidance by 73 organisations, only 40% gave specific consideration to using surrogates.[33] Such variation across HTA bodies yields heterogenous conclusions about the relevance of the same putative surrogate endpoints across different settings.[34]

Evaluating surrogate endpoints

Methodologists stress that evidence at the individual patient level alone is insufficient to evaluate surrogate endpoints especially when such evidence is obtained from a single trial.[35] This is because the observed surrogate-to-clinical outcome relationship for one drug may not hold for another, as it depends on the treatment's mechanism of action.[35] For example, progression-free survival was previously shown to be a good surrogate for overall survival in advanced colorectal cancer based on evidence from trials of traditional chemotherapy.[36] However, Ciani et al. recently observed a weaker relationship between these endpoints in this setting for modern therapies with different mechanisms of action.[37]

Meta-analysis, which combines data from a number of randomised trials, is more appropriate for evaluating the association between the treatment effects on the candidate surrogate endpoint and on the final patient-relevant clinical outcome.[38] There is growing methodological consensus for using bivariate meta-analysis methods to evaluate the surrogate-to-final outcome relationships. [39–44] These methods take into account not only the correlation between the treatment effects (quantifying the surrogate relationship), but also uncertainty around this relationship, which is crucial for decision-making.[44,45]

Table 1 lists selected examples of candidate surrogate endpoints evaluated using meta-analysis methods with authors' conclusions regarding the strength of the surrogate relationship. It is perhaps not surprising that bevacizumab's initial effect on progression-free survival never translated to prolonged survival for patients with metastatic breast cancer following FDA's accelerated approval, as an earlier meta-analysis concluded that progression-free survival was not a good surrogate for overall survival in this setting.[36]

A potential problem when evaluating surrogate endpoints is the limited amount of available randomised trial data in some areas, e.g., for drugs targeting genetic biomarkers in small patient populations. In such cases, novel bivariate network meta-analysis methods , [46] or hierarchical models,[47] allow for using readily available data on similar drugs or drug classes. These advanced methods are highlighted in reports prepared by the NICE Decision Support Unit.[44,45]

Way forward

Regulators should be more selective in their use of surrogate endpoints. Surrogate endpoints are not useful – and should not be used – when a drug's effect on the final clinical outcome can be observed within a relatively short time frame, e.g., in acute conditions.[48] Hence, using surrogate endpoints should be reserved for 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 patient-relevant clinical outcome can be assessed.[49] Even in such cases, regulators have other tools at their disposal to ensure patients who have exhausted all available treatment options can receive investigational treatments before regulatory approval. Such "expanded access" programs can bridge the access gap while evidence on patient-relevant endpoints accrues before regulatory approval.

When using surrogate endpoints is justified in selected chronic disease settings, regulators should consider the strength of available evidence on how well surrogates predict clinical benefit. The recent US accelerated approval of aducanumab for the treatment of Alzheimer's disease demonstrates why this is essential. FDA's decision was controversial in part because amyloid level changes had little to no effect on cognitive change in an earlier meta-analysis of randomized controlled trials.[1] Thus, it is still debatable whether the reduction in amyloid levels is an acceptable surrogate for cognition on the basis of current best evidence.

In the absence of regulatory guidance, there are promising signs that HTA bodies are increasingly raising the bar for using surrogate endpoints. For example, NICE has recently proposed changes to its HTA methods to strengthen the evidence requirements for the use of surrogate endpoints, while still allowing flexibility when desired evidence is not available.[50,51] 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.[52]

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.[44,46] If recommended meta-analysis methods are used, 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.

Raising the bar for using surrogate endpoints by regulators and HTA bodies may increase the cost and duration of drug development. 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.[53,54]

Greater involvement of patients (and organisations representing patients) in regulatory and HTA processes is also essential to ensure that the conditions for accepting surrogate endpoints for decision-making are adequately met. When using such endpoints is justified, patients can help ensure

that uncertainty related to surrogates is explicitly presented and taken into account. Patient input can also help guide regulatory and HTA decisions regarding the appropriate use of surrogate endpoints.

Key messages

- Surrogate endpoints are widely used by regulators to expedite the approval of new drugs, but most surrogate endpoints are not shown to be reliable predictors of outcomes that matter most to patients.
- Regulators should only accept surrogate endpoints when generating data on clinical outcomes is not attainable.
- When directly measuring drug effects on patient-relevant clinical outcomes would require trials of very substantial duration, regulators and health technology assessment bodies should systematically evaluate the appropriateness of surrogate endpoints using up to date meta-analysis methods.

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.

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 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 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

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

6 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

7 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

8 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

9 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

10 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

11 Carpenter D, Kesselheim AS, Joffe S. Reputation and precedent in the bevacizumab decision. *N Engl J Med* 2011;**365**:e3. doi:10.1056/NEJMp1107201

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 Powles T, Durán I, van der Heijden MS, *et al.* Atezolizumab versus chemotherapy in patients with platinum-treated locally advanced or metastatic urothelial carcinoma (IMvigor211): a multicentre, open-label, phase 3 randomised controlled trial. *Lancet* 2018;**391**:748–57. doi:10.1016/S0140-6736(17)33297-X

15 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

16 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

17 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

18 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

19 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

20 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

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. Assessment of the Clinical Benefit of Cancer Drugs Receiving Accelerated Approval. *JAMA Intern Med* 2019;**179**:906–13. doi:10.1001/jamainternmed.2019.0462

23 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.

24 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

25 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

26 US Food & Drug Administration. Table of Surrogate Endpoints That Were the Basis of Drug Approval or Licensure. https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure (accessed 19 Aug 2021).

27 Kim C, Prasad V. Strength of Validation for Surrogate End Points Used in the US Food and Drug Administration's Approval of Oncology Drugs. *Mayo Clin Proc* Published Online First: 10 May 2016. doi:10.1016/j.mayocp.2016.02.012

28 Haslam A, Hey SP, Gill J, *et al.* A systematic review of trial-level meta-analyses measuring the strength of association between surrogate end-points and overall survival in oncology. *Eur J Cancer* 2019;**106**:196–211. doi:10.1016/j.ejca.2018.11.012

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 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)**.

32 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

33 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

34 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

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

36 Burzykowski T, Buyse M, Piccart-Gebhart MJ, *et al.* Evaluation of tumor response, disease control, progression-free survival, and time to progression as potential surrogate end points in metastatic breast cancer. *J Clin Oncol* 2008;**26**:1987–92. doi:10.1200/JCO.2007.10.8407

37 Ciani O, Buyse M, Garside R, *et al.* Meta-analyses of randomized controlled trials show suboptimal validity of surrogate outcomes for overall survival in advanced colorectal cancer. *J Clin Epidemiol* 2015;**68**:833–42. doi:10.1016/j.jclinepi.2015.02.016

38 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

39 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

40 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

41 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

42 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

43 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

44 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. 2019. http://nicedsu.org.uk/wp-content/uploads/2020/10/TSD-20-mvmeta-final.pdf

45 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

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 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).

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

50 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

51 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

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

53 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

54 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

55 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

56 Cholesterol Treatment Trialists' (CTT) Collaborators. 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

57 Petrelli F, Borgonovo K, Cabiddu M, *et al.* Pathologic complete response and disease-free survival are not surrogate endpoints for 5-year survival in rectal cancer: an analysis of 22 randomized trials. *J Gastrointest Oncol* 2017;**8**:39–48. doi:10.21037/jgo.2016.11.03

58 Abdel-Rahman O. Surrogate end points for overall survival in trials of PD-(L)1 inhibitors for urinary cancers: a systematic review. *Immunotherapy* 2018;**10**:139–48. doi:10.2217/imt-2017-0115

59 Harshman LC, Xie W, Moreira RB, *et al.* Evaluation of disease-free survival as an intermediate metric of overall survival in patients with localized renal cell carcinoma: A trial-level meta-analysis. *Cancer* 2018;**124**:925–33. doi:10.1002/cncr.31154

60 Xie W, Regan MM, Buyse M, *et al.* Event-Free Survival, a Prostate-Specific Antigen-Based Composite End Point, Is Not a Surrogate for Overall Survival in Men With Localized Prostate Cancer Treated With Radiation. *J Clin Oncol* 2020;**38**:3032–41. doi:10.1200/JCO.19.03114

61 Hughes M, Daniels M, Fischl M, *et al.* CD4 cell count as a surrogate endpoint in HIV clinical trials: a meta-analysis of studies of the AIDS Clinical Trials Group. *AIDS* 1998;**12**:1823–32.

Table 1: Examples of candidate surrogate endpoints evaluated using meta-analysis and authors' conclusions regarding the strength of the surrogate relationship

Disease area	Candidate surrogate endpoint	Final clinical outcome	Strength of the surrogate relationship, as reported by study authors
Gastric cancer [8]	Disease-free survival	Overall survival	"Disease-free survival is an acceptable surrogate for overall survival in trials of cytotoxic agents for gastric cancer in the adjuvant setting"
Multiple sclerosis [55]	Relapse rate	Expanded Disability Status Scale (EDSS) worsening	"support the use of commonly used surrogate markers of expanded disability status scale worsening as endpoints in multiple sclerosis clinical trials"
Immunoglobulin A nephropathy [56]	Change in proteinuria	Doubling of serum creatinine level, end-stage renal disease, or death	"supporting the use of an early reduction in proteinuria as a surrogate endpoint for clinical end points in Immunoglobulin A nephropathy in selected settings"
Cardiovascular disease	Low-density lipoprotein	Major coronary events	"an approximately linear relationship between the absolute reductions in low-density lipoprotein cholesterol achieved in these trials and the proportional reductions in the incidence of coronary and other major vascular events"
Advanced colorectal cancer in traditional chemotherapy trials [56]	Progression-free survival	Overall survival	"PFS is an acceptable surrogate for OS in advanced colorectal cancer"
Advanced colorectal cancer in modern trials [37]	Progression-free survival	Overall survival	"none of the end points were found to achieve the level of evidence (i.e., mean R^2_{trial} > 0.60) that has been set to select high or excellent correlation levels by common surrogate evaluation tools"
Metastatic breast cancer [36]	Tumour response, disease control, progression-free survival, and time-to-progression	Overall survival	"no end point could be demonstrated as a good surrogate for overall survival in these trials"
Rectal cancer [57]	Pathologic complete response and disease-free survival	Overall survival	"pathologic complete response and disease-free survival are not surrogate endpoints for 5-year survival in rectal cancer"
Urinary cancer [58]	Overall response rate and progression-free survival	Overall survival	"overall response rate and progression-free survival are not reliable surrogate end points for median overall survival in trials of PD-(L)1 inhibitor therapy for urinary cancers"

Renal cell carcinoma [59]	Disease-free survival	Overall survival	“there was no strong correlation noted between 5-year disease-free survival and 5-year overall survival rates or between treatment effects on these endpoints.”
Prostate cancer [60]	Event-free survival	Overall survival	“event-free survival is a weak surrogate for overall survival and is not suitable for use as an intermediate clinical end point to substitute for overall survival”
HIV infection [61]	CD4 count	AIDS or death	“CD4 cell count is a weak surrogate endpoint”
Alzheimer’s disease [1]	Amyloid levels	Cognitive decline	“reducing amyloid levels with drug treatment has, at most, a small effect on cognition”

474
475
476