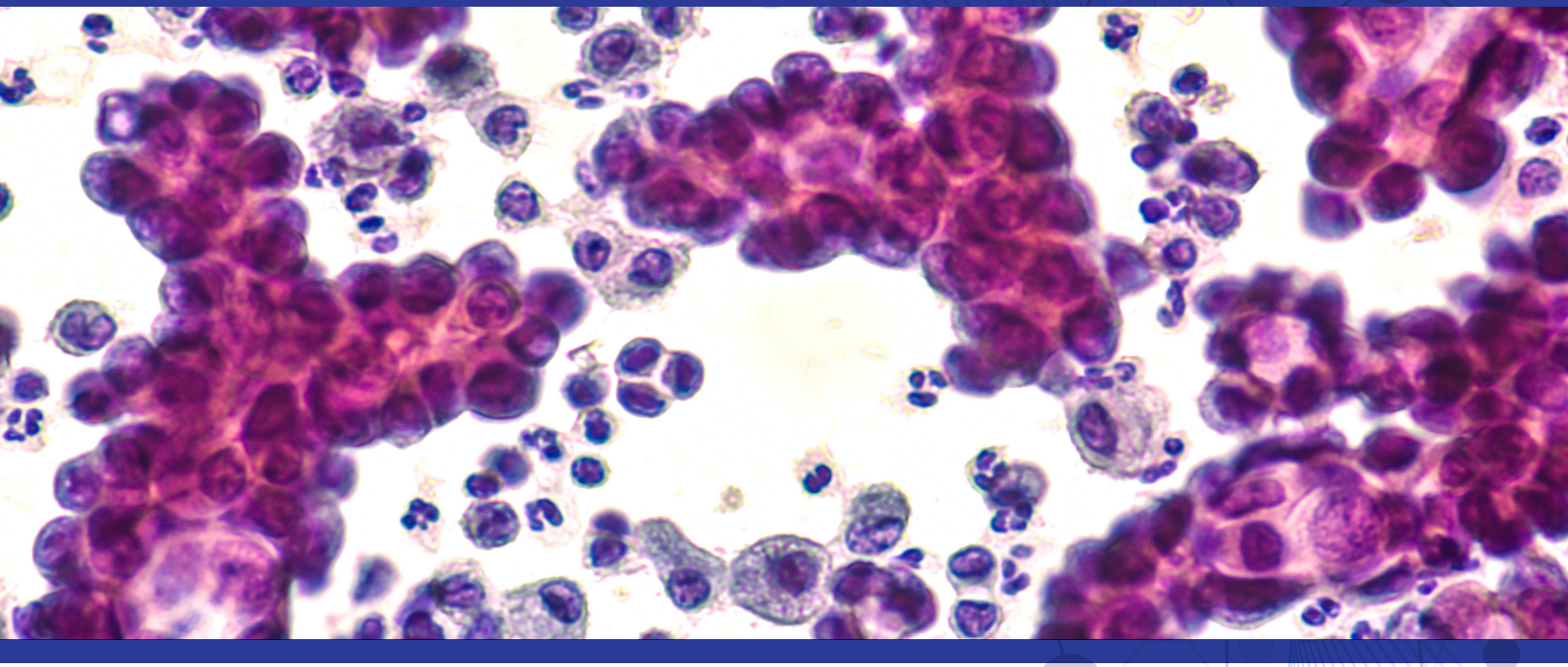
CERTARA

Surrogate Endpoints Under Attack – Is It Still Worth Performing Surrogacy Validation? Lessons From NSCLC

White Paper summarizing the ISPOR US 2022 Issue Panel #223



By Dr. Silvia Paddock, Dr. Dalia Dawoud, Dr. Billy Amzal, and Dr. Jeff Allen

Surrogate Endpoints Under Attack – Is It Still Worth Performing Surrogacy Validation? Lessons From NSCLC

This white paper summarizes the ISPOR US 2022 Issue Panel #223

Executive Summary

The past 50 years have witnessed a vigorous debate around the use of surrogate (aka intermediate) endpoints in oncology and other disease areas.

- Reliance on inadequate surrogate endpoints has led to waste of resources and, at its worst, even harm to patients
- At the same time, well-characterized surrogate endpoints have been shown to be vital when equipoise cannot be maintained in clinical trials otherwise, and for providing timely access to beneficial treatments years sooner than awaiting long-term clinical outcomes

This panel brought together experts from the fields of health technology assessment (HTA), statistical modeling, and patient advocacy to contribute to the important ongoing dialogue among stakeholders from different perspectives.

NICE methods update for surrogate endpoint validation (Dr. Dalia Dawoud, Senior Scientific Adviser, NICE, UK):

- The new National Institute for Health and Care Excellence (NICE) methods manual requires more explicit evidence when using surrogate endpoints in submissions
- A hierarchy of levels of evidence is introduced (top to bottom):
    - o Evidence from randomized clinical trials (RCTs) demonstrating correlation of the treatment effect on the surrogate endpoint with the treatment effect on the target endpoint
    - o Observational or real-world evidence (RWE) showing consistent association between the surrogate and target measures
    - o Biological plausibility of the surrogate/target- endpoint relationship
- It is recommended to use bivariate meta-analytic approaches to determine the correlation of the treatment effect on the surrogate endpoint with the treatment effect on the target endpoint
- The uncertainty in the final endpoint prediction should be quantified and accounted for in the economic models

Statistical methodology update (Dr. Billy Amzal, CEO, Quinten Health, France)

- The choice of the adequate method and statistical model to assess surrogacy is critical to best capture heterogeneity and potential non-linearities and hence reduce bias and uncertainty when predicting the final endpoint based on the surrogate/intermediate endpoint

- For non-small cell lung cancer (NSCLC) specifically, analyses of a large clinical dataset showed that use of joint (aka bivariate) models, recommended in the new NICE guidelines, makes most use of the available information and can improve of the surrogacy estimates substantially
- RWE may provide additional information where clinical trial evidence is unavailable, too sparse or when real-world outcomes matter more

The value of surrogates for oncology patients (Dr. Jeff Allen, President and CEO, Friends of Cancer Research, USA)

- The accelerated approval pathway is an important route for treating severe conditions to reach patients faster, and these approvals rely on surrogate endpoints along with the commitment to develop confirmatory evidence regarding long-term outcomes
- One needs to expect that it will require additional time to evaluate final clinical outcomes; however, diligent follow-up on studies to confirm the initial effect on surrogates is important and follow-on studies should be initiated and completed in a timely manner to demonstrate actual benefits for the patients
- More frequently than not, accelerated approvals based on surrogate endpoints in oncology have subsequently confirmed clinical benefit with only approximately 10% of accelerated approvals being withdrawn to date

Conclusions and outlook

The panel agreed that surrogacy validation in NSCLC is an as timely endeavor as it has ever been and that alignment on requirements and optimal methodology (such as using the bivariate/joint modelling approach as recommended by NICE) provides an important roadmap for early dialogue between manufacturers, regulators, and payers.

- Active engagement of all stakeholders is necessary to align on a consistent and rational methodology that allows all sides to weight benefits against risks inherent to the use of surrogates
- Policy requirements to ensure safe and yet timely access to treatments include the timely completion of confirmatory studies and transparent and streamlined withdrawal processes

Inviting patient representatives to panels such as this one was highlighted as an important step toward making sure that patient benefit is always front and center in considerations of surrogate endpoints.

Objective

This white paper summarizes an experts' discussion on surrogate endpoints, held at the issue panel #223 "Surrogate Endpoints Under Attack – Is It Still Worth Performing Surrogacy Validation? Lessons From NSCLC" at ISPOR 2022 in Washington D.C., US, with a focus on how to best tackle analytical issues in surrogate endpoint validation.

Introduction

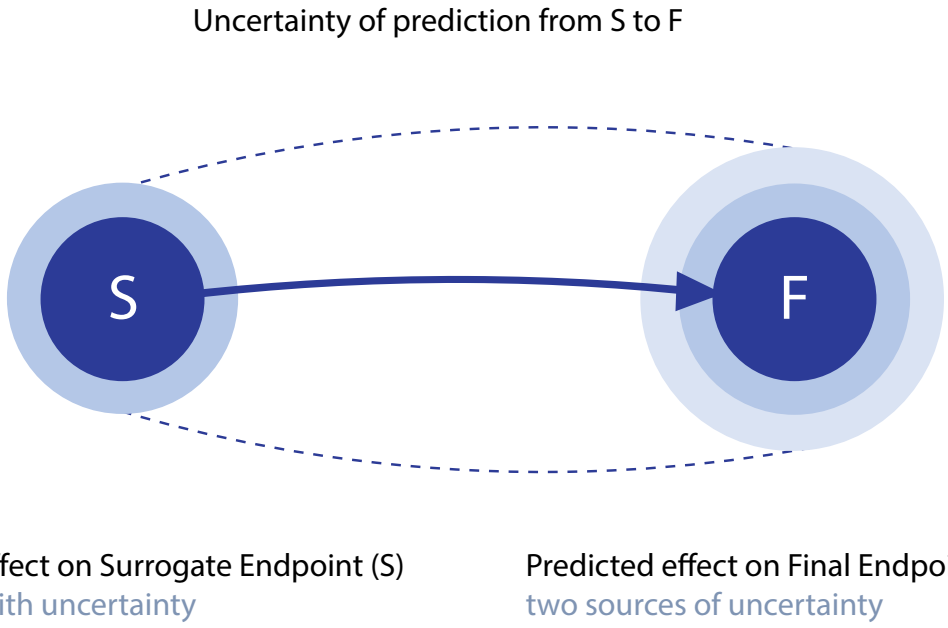
Dr. Silvia Paddock, Senior Manager, PwC Switzerland

In her introduction to the topic of surrogate endpoints and their validation, Dr. Paddock highlighted the importance of "uncertainty" for the current debate and explained that in general, the prediction of the final endpoint based on the surrogate endpoint includes two sources of uncertainty: one source of uncertainty from measuring the surrogate endpoint, and another source of uncertainty from predicting the final endpoint based on the surrogate endpoint, illustrated in Figure 1.

Dr. Paddock noted that this uncertainty necessarily implies the risk of being wrong about surrogacy, and that failures on either side of the spectrum (i.e., using surrogates that turned out to be harmful or not using surrogates when equipoise was violated in trials) have led to the debate around surrogates becoming quite emotional and the pendulum of their acceptability swinging widely between "absolutely vital" and "completely unacceptable". She expressed hope that panels such as this one could contribute to a more "stable equilibrium" of surrogate use based on commonly agreed best practices and shared evaluation of the benefit/risk ratio for patients.

Figure 1.

Sources of uncertainty when predicting the final endpoint based on the surrogate endpoint



Surrogate endpoint validation in NSCLC

Dr. Paddock elaborated on attempts to validate progression-free survival (PFS) or overall response rate (ORR) as surrogate endpoints for overall survival (OS) in non-small cell lung cancer (NSCLC) that have been difficult in the past[1,2], which could partially be explained by cross-over from the control to the treatment arm after progression.[2] Other potential issues hampering surrogate validation in NSCLC are pseudoprogression[3] (a phenomenon in which an initial increase in tumor size is observed or new lesions appear, followed by a decrease in tumor burden), which is observed in some patients on immunotherapies, and multiple later lines of treatment "diluting" the effect of the initial treatment. At the same time, the biological rationale that PFS is a potential surrogate for OS is strong in NSCLC, since patients die from cancer in the lung[4], thus, a thorough and comprehensive assessment of surrogacy, using advanced analytic methods that reduce biases compared to prior analyses, is warranted.

Experts' opinions on surrogate endpoints

In the following, different perspectives of the panelists on various aspects of surrogacy validation, hurdles for using of surrogate endpoints, and possible solutions to overcome those hurdles are presented.

HTA/PAYERS' PERSPECTIVE

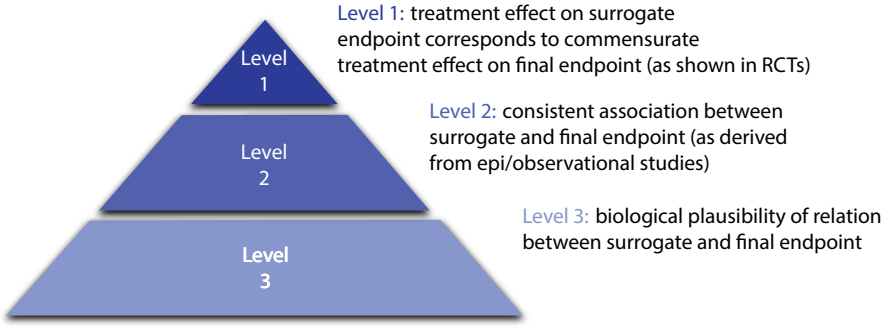
Dr. Dalia Dawoud, Senior Scientific Adviser, National Institute for Health and Care Excellence (NICE), UK

Dr. Dawoud pointed out that HTA agencies increasingly reviewing submissions containing early evidence and immature data. In particular, the evidence presented to support the validity of the relationship between the surrogate endpoints and outcomes required for decision-making is often limited. NICE therefore recently updated its methods guideline[5] on surrogacy validation after reviewing (i) current methods used within NICE, (ii) those used in other HTA organizations, and (iii) key literature on the topic including a commissioned report by NICE Decision Support Unit[6]. Dr. Dawoud explained that the updated guideline calls for more explicit evidence in the submissions and provides methodological guidance on how to prepare this evidence. Three key changes in the methods update are laid out in the following.

First, three levels of evidence can be considered[7] (Figure 2), with the most important level (Level 1) being the correspondence of treatment effects on the surrogate endpoint to the treatment effects on the final endpoint as shown in RCTs. Robust meta-analytic methods exist to conduct such analyses, for example joint modelling (aka bi-variate meta-analysis) of the surrogate and the final endpoint[6,8] are recommended.

Figure 2.

Levels of evidence to be considered in surrogate analyses



Second, the evidence presented should be specific to the population, disease area and technology of interest. Deviations from this rule, e.g. for first-in-class treatments, need thorough justification, and using meta-analytic methods that allow borrowing of information from sufficiently similar treatment classes, populations, and treatment settings is recommended.[8]

Third, the uncertainty on predicting the final outcome from the surrogate outcome needs to be quantified, to enable its incorporation into the economic models through probabilistic sensitivity analyses.

METHODOLOGICAL PERSPECTIVE

Dr. Billy Amzal, CEO, Quinten Health, France

Dr. Amzal explained that in addition to the challenges of surrogacy validation mentioned above that are more specific to oncology/NSCLC, surrogacy validation is often performed with aggregated data on published study results which limits the statistical methods to adjust for confounders.

If the functional relationship between the surrogate and the final endpoint is mis-specified, the available data may not be used efficiently, and popular methods such as meta-regression disregard that also the surrogate endpoint is measured with error. Therefore, choosing the appropriate statistical method that can avoid bias and reduces the uncertainty of the prediction is critical. For example, covariate adjustment, even if performed on the aggregate level, may reduce bias and increase precision.

The case of NSCLC

Dr. Amzal presented results from an exploratory analysis that was performed on a large NSCLC database including aggregate data of more than 1000 clinical trials published before the end of June 2021. The aim of this data exploration was to evaluate:

- How strength of surrogacy or uncertainty may be impacted by different model parameterization, and to benchmark different statistical models,
- How surrogacy may vary across patient subgroups.

Both trial-level (hazard ratios, HRs) and arm-level (median survival times) results were used to assess PFS and ORR as surrogate endpoints for OS.

This exploratory analysis revealed that surrogacy strength depended on the model parameterization and how well nonlinearities and collinearities are captured. For example, surrogacy strength substantially improved when using a joint/bi-variate meta-analytic model[8] for the surrogate and the final endpoint, which can be seen in Table 1: the adjusted R^2 (correlation coefficient) of the joint model increased to 0.47 as compared to the one of the standard meta-regression model. In addition, while joint modelling does not impact mean predictions, it led to narrower prediction intervals compared to those based on the standard meta-regression model, see the last column of Table 1.

Table 1.

Comparison of adjusted R^2 and predicted HRs from the standard meta-regression model

Model	Adjusted R^2 [95% CI]	HR PFS	Predicted HR OS with 95% PrI
Standard meta-regression model	0.26 [0.11, 0.41]	0.4	0.69 [0.54, 0.88]
		0.6	0.80 [0.64, 1.00]
		0.7	0.84 [0.68, 1.05]
Joint/bi-variate model	0.47 [0.27, 0.64]	0.4	0.70 [0.59, 0.84]
		0.6	0.80 [0.67, 0.96]
		0.7	0.85 [0.71, 1.01]

CI: confidence interval, PrI: prediction interval

The exploratory analysis also showed that surrogacy strength varied across subgroups and could thus be improved by adjusting for relevant covariates. For example, surrogacy for PFS vs. OS resulted in R^2 = 0.501 for trials with high percentage of patients who never smoked as compared to R^2 = 0.339 overall. Likewise, and as expected, surrogacy depends on drug class: we found for example a very weak surrogacy in immunotherapy-treated subgroups (R^2=0.007) but a very strong surrogacy in chemo-treated subgroups (R^2=0.843).

In summary, Dr. Amzal concluded that this exploratory analysis for NSCLC thus negated the simplistic message that "surrogacy is weak" in NSCLC as previously published. It also indicated a clear correlation between PFS gain and OS gain: progression rates reduction of X% translated into mortality rates reduction of X/2%, and longer median PFS seemed to be associated with longer median OS. The study confirmed that the surrogacy question needs appropriate methods (e.g., using non-linear and/or joint models) to be robust and fair.

Surrogacy validation using real-world evidence

Dr. Amzal further elaborated on the potential inclusion of real-world data (RWD) in surrogacy validation analyses: if (aggregated or individual-level) data are sparse for the drug class and/or population studied, RWD may be used to enhance and generalize surrogacy. Such an analysis requires mapping of relevant variables from RCTs to the real-world, and it needs to account for drivers of effectiveness. Though no method to integrate RWD into surrogacy validation is recommended yet by NICE, research is active in the field and a growing number of pilots and cases have been explored, some submitted to HTA agencies. It required appropriate methodologies, such as the use of pharmaco-epidemiological methods to control potential biases or the use of Bayesian bivariate random-effects meta-analysis as proposed by Wheaton et al.[9]

PATIENT AND POLICY PERSPECTIVE

Dr. Jeff Allen, President and CEO, Friends of Cancer Research (FOCR), USA

Accelerated approval has enabled patients access to new medicines to treat serious illnesses with unmet medical need years sooner than would have been possible without this pathway. Dr. Allen explained that yet, besides withdrawn indications amounting to 10.1% among all accelerated approvals, a large portion of indications with pending action have not yet been converted to full approval, as can be seen in Table 2. He also pointed out that the average time that had passed since these were approved was still shorter than the average gain that previous approvals had brought in terms of earlier access. This means that the data may not be mature yet for a substantial share of the pending molecules. This finding calls for patience

when waiting for final results but also timely initiation and completion of confirmatory studies and a transparent and streamlined withdrawal process at the regulatory level.

Dr. Allen highlighted that oncology has benefitted from accelerated approval due to quantifiable and standardized disease measures. This has been supported by rapidly evolving science leading to an improving understanding of disease biology and advanced drug design. Most importantly the patient benefit is becoming clear – a recent study demonstrates a reduction in the 5-year mortality of NSCLC due to the availability of targeted therapies[12] many of which were approved through accelerated approval.

Dr. Allen further reviewed recent work by his organization to validate circulating tumor DNA as a potential surrogate endpoint for future use in cancer research and showed recent data from a study that pooled results from several RCTs to investigate the potential utility of this marker. He noted the importance of such rigorous validations to ensure that patients truly benefit from adequate use of surrogate endpoints.

Table 2.

Overview of accelerated approvals across all indications and for oncology indications

Accelerated Approval Indications	Total with Accelerated Approval	Total Converted to Full Approval	Total Withdrawn	Total Pending Action
All[10]	278	139 (50.0%)	28 (10.1%)	111 (39.9%)
All Approved Over 5-Years	66	34 (51.5%)	14 (21.2%)	18 (27.3%)
Oncology	190	84 (44.2%)	17 (9.0%)	89 (46.8%)
Oncology Products Over 5-Years[11]	33	19 (57.6%)	5 (15.2%)	9 (27.3%)

Conclusions

The panel concluded that surrogacy validation in NSCLC was an as timely subject as ever, and that alignment on requirements and optimal methodology, discussed in the current panel from multiple perspectives, provide an important roadmap for early dialogue between manufacturers, regulators, HTA agencies and payers.

Previous challenges hampering surrogacy validation may be alleviated by selecting appropriate analytical methods such as joint/bi-variate meta-analytic models for the surrogate and final endpoint that improve surrogacy strength in terms of the R^2 and reduce uncertainty of the final endpoint prediction. For NSCLC, contrary to the pessimistic tenor of recent surrogacy validation attempts, a clear correlation between PFS gain and OS gain could be observed based on RCT evidence with optimized methodology. Proper accounting for all uncertainties in the calculations in the future will help HTA agencies to better accommodate early evidence from surrogate endpoints. Because no methodology will ever fully eliminate all uncertainties when using surrogates, an intense dialogue is necessary to ensure that the benefit/risk balance reflects the patient perspective. Surrogate use needs to be planned early during trial design, and diligent follow-on studies are necessary to ensure that patients indeed benefit from earlier access to targeted therapies.

About the Authors

Dr. Silvia Paddock is a Senior Manager in the Health Analytics division of PwC Switzerland. A neuroscientist by training, she has provided consulting services for more than 10 years in the areas of evidence synthesis, indirect treatment comparisons, economic modeling, market access, and policy. Her current main focus areas are the agile development of decision support tools in learning healthcare systems and the use of federated learning solutions for the integration of insights from diverse data sources.

Dr. Dalia Dawoud is Senior Scientific Adviser at the National Institute for Health and Care Excellence (NICE). She holds MSc in Economic Evaluation in Health Care from City University London and PhD in pharmaceutical policy and economics from King's College London. She has long experience in using economic evaluation in clinical guidelines development and health technology assessment (HTA), gained through working on NICE Clinical Guidelines as well as technology appraisals. Dr. Dawoud's research interests are focused on the advanced methods of evidence synthesis and use in economic models and the use of real-world evidence to inform drug development and health care decision making. Dalia currently has overall responsibility of overseeing the delivery of NICE allocated tasks on a portfolio of IMI and Horizon 2020 funded research projects including EHDEN and HTx. She is widely published in the field of pharmaceutical policy and pharmacoeconomics. She also serves as Associate Editor for ISPOR journal Value in Health and as Associate Editor for Pharmacoeconomics and Outcomes Research for Elsevier's journal Research in Social and Administrative Pharmacy. Dalia also holds adjunct position as Associate Professor at the Faculty of Pharmacy, Cairo University.

Dr. Billy Amzal graduated from Ecole Polytechnique, from AgroParisTech and holds a PhD in Decision Mathematics from Paris-Dauphine University which was awarded by the International Society of Bayesian Analysis and by the International Biometrics Society. Over the last 20 years, Billy has developed predictive analytics tools to inform and support strategic decision making in healthcare. Prior to joining Quinten Health as CEO, Billy developed predictive and impactful tools within big pharma companies and public health agencies. He then led consultancy teams at Certara focusing on RWD and decision analytics as senior VP, acted as statistical expert for public Health Authorities (EFSA, ANSES, WHO), and taught pharmaco-economics at CNAM university in Paris. Billy authored more than 100 scientific publications in international journals. He conceived and led hundreds of disease modeling projects supported 100s of HEOR model submissions to HTAs and real-world database studies. He created Quinten Health, the leading company in RW simulation and disease modeling.

Dr. Jeff Allen serves as the President and CEO of Friends of Cancer Research (Friends). For over 25 years, Friends has created unique scientific partnerships, accelerated policy change, and supported groundbreaking research to deliver new therapies to patients quickly and safely. As a key thought leader on issues related to the U.S. Food and Drug Administration, healthcare, and regulatory policy, he is regularly published in prestigious medical journals and policy publications and has contributed his expertise to the legislative process on multiple occasions.

CERTARA

Acknowledgements

This ISPOR panel and the NSCLC case study analyses received financial support from Amgen. We gratefully acknowledge the support of Dr. Noemi Hummel (Associate Director, Decision Analytics & Modeling, Certara) in writing this white paper, and we thank Dr. Shuai Fu (Senior Manager, Decision Analytics & Modeling, Certara) and Agnieszka Kopiec (Manager, Decision Analytics & Modeling, Certara) for performing the NSCLC case study analyses.

References

1. Blumenthal GM, Karuri SW, Zhang H, et al. Overall response rate, progression-free survival, and overall survival with targeted and standard therapies in advanced non-small-cell lung cancer: US Food and Drug Administration trial-level and patient-level analyses. J Clin Oncol Off J Am Soc Clin Oncol. 2015;33(9):1008-1014. doi:10.1200/JCO.2014.59.0489

2. Clarke JM, Wang X, Ready NE. Surrogate clinical endpoints to predict overall survival in non-small cell lung cancer trials-are we in a new era? Transl Lung Cancer Res. 2015;4(6):804-808. doi:10.3978/j.issn.2218-6751.2015.05.03

3. Ma Y, Wang Q, Dong Q, Zhan L, Zhang J. How to differentiate pseudoprogression from true progression in cancer patients treated with immunotherapy. Am J Cancer Res. 2019;9(8):1546-1553.

4. Nichols L, Saunders R, Knollmann FD. Causes of death of patients with lung cancer. Arch Pathol Lab Med. 2012;136(12):1552-1557. doi:10.5858/arpa.2011-0521-OA

5. NICE. NICE Health Technology Evaluations: The Manual.; 2022. Accessed May 27, 2022. https://www.nice.org.uk/process/pmg36/resources/nice-health-technology-evaluations-the-manual-pdf-72286779244741

6. Welton NJ, Phillippo DM, Owen R, et al. CHTE2020 Sources and Synthesis of Evidence; Update to Evidence Synthesis Methods.; 2020. https://www.sheffield.ac.uk/sites/default/files/2022-02/CHTE-2020_final_20April2020_final.pdf

7. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. Value Health J Int Soc Pharmacoeconomics Outcomes Res. 2017;20(3):487-495. doi:10.1016/j.jval.2016.10.011

8. 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. Stat Med. 2020;39(8):1103-1124. doi:10.1002/sim.8465

9. Wheaton L, Papanikos A, Thomas A, Bujkiewicz S. Using Bayesian Evidence Synthesis Methods to Incorporate Real World Evidence in Surrogate Endpoint Evaluation. Published online December 16, 2021. Accessed May 22, 2022. arXiv:2112.08948v1

10. Kaltenboeck A, Mehlman A, Stephen Pearson. Strengthening the Accelerated Approval Pathway: An Analysis of Potential Policy Reforms and Their Impact on Uncertainty, Access, Innovation, and Costs.; 2021. https://icer.org/wp-content/uploads/2021/04/Strengthening-the-Accelerated-Approval-Pathway-_-ICER-White-Paper-_-April-2021.pdf

11. Friends of Cancer Research. Drug Development Dashboard. Accessed February 9, 2022. https://friendsofcancerresearch.org/drug-development-dashboard/

12. Howlader N, Forjaz G, Mooradian MJ, et al. The Effect of Advances in Lung-Cancer Treatment on Population Mortality. N Engl J Med. 2020;383(7):640-649. doi:10.1056/NEJMoa1916623

About Certara

Certara accelerates medicines using proprietary biosimulation software, technology and services to transform traditional drug discovery and development. Its clients include more than 2,000 biopharmaceutical companies, academic institutions, and regulatory agencies across 62 countries.

For more information visit www.certara.com or email sales@certara.com.

WP July 2020 v1 070920