Cancer Management and Research

Elevate Series

Open Access Full Text Article

PERSPECTIVES

UK Stakeholder Perspectives on Surrogate Endpoints in Cancer, and the Potential for UK Real-World Datasets to Validate Their Use in Decision-Making

David Baldwin[1], Jonathan Carmichael[2], Gordon Cook[3], Neal Navani[4,5], James Peach[6], Ruth Slater[7], Pete Wheatstone[8], Julia Wilkins[7], Nicola Allen-Delingpole[9], Cicely EP Kerr[10], Khalid Siddiqui[11]

[1]Department of Respiratory Medicine, Nottingham University Hospitals NHS Trust and the University of Nottingham, Nottingham, UK; [2]Department of Oncology, The National Institute for Health Research Leeds In Vitro Diagnostics Co-Operative (NIHR Leeds MIC), Leeds, UK; [3]Cancer Research UK Trials Unit, LICTR, University of Leeds & NIHR (Leeds) IVD MIC, Leeds, UK; [4]Lungs for Living Research Centre, UCL Respiratory, University College London, London, UK; [5]Department of Thoracic Medicine, University College London Hospital, London, UK; [6]Human Centric Drug Discovery, Wood Centre for Innovation, Oxford, UK; [7]Health Economics Unit, London, UK; [8]Patient and Public Involvement and Engagement Group, DATA-CAN, London, UK; [9]The Association of the British Pharmaceutical Industry, London, UK; [10]Research Data Scotland, Bayes Centre, Edinburgh, UK; [11]Janssen-Cilag Ltd, High Wycombe, UK

Correspondence: James Peach, Human Centric Drug Discovery, Wood Centre for Innovation, Quarry Road, Oxford, OX3 8SB, UK, Email james@humancentricdd.com

Abstract: Duration of overall survival in patients with cancer has lengthened due to earlier detection and improved treatments. However, these improvements have created challenges in assessing the impact of newer treatments, particularly those used early in the treatment pathway. As overall survival remains most decision-makers' preferred primary endpoint, therapeutic innovations may take a long time to be introduced into clinical practice. Moreover, it is difficult to extrapolate findings to heterogeneous populations and address the concerns of patients wishing to evaluate everyday quality and extension of life. There is growing interest in the use of surrogate or interim endpoints to demonstrate robust treatment effects sooner than is possible with measurement of overall survival. It is hoped that they could speed up patients' access to new drugs, combinations, and sequences, and inform treatment decision-making. However, while surrogate endpoints have been used by regulators for drug approvals, this has occurred on a case-by-case basis. Evidence standards are yet to be clearly defined for acceptability in health technology appraisals or to shape clinical practice. This article considers the relevance of the use of surrogate endpoints in cancer in the UK context, and explores whether collection and analysis of real-world UK data and evidence might contribute to validation.

Keywords: cancer, surrogate endpoints, real-world data, multiple myeloma, lung cancer, quality of life

Introduction

Advances in understanding the mechanisms underlying cancer have driven the development of innovative therapeutic approaches, and half of patients diagnosed with cancer in the UK are now predicted to survive for 10 years or more.[1] Despite this progress, the timely evaluation of new cancer treatments remains challenging because overall survival (OS) continues to be the regulators' and health technology assessment (HTA) agencies' preferred endpoint. This endpoint offers many positives: OS and quality of life (QoL) are the two endpoints most important to patients; it is a hard endpoint that is relevant in all cancer studies; and strong results for novel drugs are reflected by early signals of OS benefit. Nevertheless, achievement of median OS relies on large cohorts and long-term follow-up. Additionally, in early-stage cancers and malignancies that progress slowly or have good long-term prognoses, which often require multiple therapeutic agents or combination regimens, it can be challenging to determine the true impact of individual therapies using OS alone.[2] Interpretation of the results can be confounded by use of other therapies – as the numbers of treatment options increase, so too do the potential combinations and sequences in which they may be used. However, the number of patients available in whom to test them remains insufficient,[3] making it difficult to clearly attribute treatment effects.[4] Additionally, diagnostic approaches (eg, liquid biopsy), advanced imaging, and data analysis techniques (eg, artificial intelligence and machine

Received: 21 September 2023
Accepted: 24 June 2024
Published: 12 July 2024

learning) have altered the ways in which safety and efficacy can be assessed,[5] and have led to alternative endpoints being used to evaluate the clinical impact of treatment alongside OS.

Randomized controlled trials (RCTs) are subject to strict eligibility criteria, which often result in high clinical trial exclusion rates that can hinder recruitment and diminish the generalizability of results.[4] Novel research approaches and the use of surrogate or interim endpoints have been applied by researchers in an attempt to answer clinical questions more quickly and affordably in a wider range of patients, albeit with variable success.[6] As a result, OS is no longer the most frequent primary endpoint in oncology RCTs.[7] Although surrogate endpoints have performed well in some cancers when assessed in patients that match RCT participants,[8] this is not ubiquitous, and there are examples in cancers, such as multiple myeloma (MM), where correlations have been poor.[9] Reliance on these early data alone and the use of surrogates as primary endpoints has, therefore, been challenged.[2,10] Thus, so far, HTA agencies have handled surrogate endpoint evidence on a case-by-case basis, but none is yet deemed sufficiently validated to be approved universally or accepted as a replacement for any conventional clinical primary endpoint. As a consequence, there is substantial variation in approaches.[11] Validation would be likely to require extensive evidence, preferably derived through meta-analysis of RCTs.[11,12] Additionally, it has been suggested that each surrogate endpoint should be validated per indication (including disease grade or stage) and per intervention.[12] Therefore, the use of surrogate endpoints so far has been assessed in terms of biomarkers expected to predict clinical benefit with reasonable likelihood and/or biological plausibility of the relationship between the surrogate and the final clinical endpoint.[11,13,14]

This Perspective paper summarizes discussions held during a feasibility assessment of the use of surrogate endpoints in cancer research and ways in which real-world data (RWD) and real-world evidence (RWE) might be used to validate such endpoints. Views were gathered from UK patient, clinical, and pharmaceutical industry representatives on the current landscape of surrogate endpoints and RWD, and the conclusions incorporate the expressed perspectives and recommendations. Two exemplar cancer types were discussed – a blood cancer (MM) and a solid tumor cancer (lung cancer [LC]) – with consideration given to themes that might be common to other cancer types.

Surrogate Marker Landscape

Regulatory Approval and Health Technology Assessments

Regulatory interest in surrogate endpoints in cancer treatment trials has been growing, particularly with the ambition of achieving expedited access to novel treatments where there is a significant unmet need.[13–17] Drug development times may be substantially shortened by the use of surrogate endpoints, which could result in decision-making being based on earlier data and allow earlier access for patients. For example, Chen et al[18] demonstrated that using surrogate endpoints for oncology drug approval was associated with a reduction in drug development time of approximately 11 months compared with an OS endpoint. However, this approach is thought to come at the cost of increased clinical uncertainty and, potentially, a shortage of information for decision-makers (eg, HTA agencies and clinicians). Therefore, data are expected to be gathered alongside those for conventional clinical trial endpoints and outcomes tested by extended follow-up to ensure they robustly correlate with meaningful endpoints for patients.[10,19] The US Food and Drug Administration (FDA) has allowed the use of surrogate endpoints in some individual drug development programs, which it lists online,[20] but still considers each case individually. The European Medicines Agency does not have a similar list, but research has shown a trend in awarding marketing authorizations for drugs in which OS data were immature and progression-free survival (PFS) was given more consideration.[21,22] The UK Medicines and Healthcare products Regulatory Agency (MHRA) has not yet issued guidance on the use of surrogate endpoints, but in January 2021, it introduced the Innovative Licensing and Access Pathway (ILAP) process, which provides a channel to discuss new approaches to evidence generation.[23] The ambition of this new licensing and access pathway was to reduce the time to market for innovative medicines. As such, ILAP presents a unique opportunity for companies to align with regulators and HTA bodies on the quality of the evidence and its utility in decision-making that is provided by a specific surrogate endpoint. So far, three drugs for cancer have been awarded an Innovation Passport via ILAP and approved by the National institute for Health and Care Excellence (NICE) and the Scottish Medicines Consortium: sotorasib for treating *KRAS* Gly12Cys mutation-positive non-small-cell LC (NSCLC), lorlatinib for NSCLC positive for *ALK* mutations, and sacituzumab govitecan for

metastatic triple-negative breast cancer. Surrogate endpoints from the pivotal trials for these drugs supported the evidence packages that informed regulatory and HTA approvals.[24]

In contrast to regulators, HTA agencies and payers have remained more conservative, and conventional survival data endpoints continue to be preferred for economic assessments. Surrogate endpoints are generally considered to provide a greater degree of uncertainty compared to OS for the extrapolation and estimation of long-term benefits to patients.[25] Furthermore, there is substantial variation between HTAs in how agencies handle data and try to validate surrogate endpoints.[26] For example, NICE considered expert opinion when deciding whether PFS was a suitable surrogate measure for the effects of brentuximab vedotin in CD30-positive Hodgkin lymphoma,[27] whereas meta-analyses of individual patient data in RCTs were used to evaluate the acceptability of pathological complete response to support pertuzumab in human *EGFR2*-positive breast cancer.[28] By contrast, the Canadian Agency for Drugs and Technologies in Health concluded that meta-analysis of pooled individual patient data was insufficient to support the validity of pathological complete response as a proxy for long-term outcomes in breast cancer.[29] The acceptability of time from randomization to progression on second-line therapy, often referred to as PFS2, has been judged differently by six national HTAs in Canada and Europe.[30]

In its health technology assessment guidance, NICE recommends "in all cases, the uncertainty associated with the relationship between the endpoints and the final outcomes should be quantified and presented" as part of a probabilistic sensitivity analysis.[31] Therefore, as part of its decision-making, NICE considers long-term effects even if complete evidence is unavailable. This approach has been evidenced in recent health technology appraisals, such as that for osimertinib where disease-free survival (DFS) was used in the approval of its use for adjuvant treatment of stage 1b to 3a NSCLC after complete tumor resection, in adults whose tumors have *EGFR* exon 19 deletions or exon 21 (Leu858Arg) substitution mutations.[32] Another example was the use of minimal (also called measurable) residual disease (MRD) data to support the approval of daratumumab in combination as an option for transplant-eligible patients with untreated MM.[33] Acknowledging the difficulty of the current situation, though, NICE is working with organizations in other countries to develop more guidance for pharmaceutical companies on the use of surrogate outcomes when analyzing cost-effectiveness.[34] The IQVIA Institute for Human Data Science noted two literature reviews, one by Cooper et al[35] and one by Pasalic et al,[36] that could indicate why payers remain skeptical of surrogate endpoints. Both showed that the predictive value of surrogate endpoints is not always reliable, meaning that payers must extrapolate and estimate true benefits to patients. This approach is believed to reduce the certainty about the economic value of treatments. One solution has been to grant conditional reimbursement, where they will accept submission of RWE after reimbursement to reduce the impact of uncertainty on future estimates. IQVIA performed a qualitative assessment with 24 US and European payers and key opinion leaders to explore their opinions about how to prioritize six surrogate endpoints (DFS, MRD, pathologic complete response, disease control rate, relapse rate, and time to response).[37] These expert stakeholders believed that, despite current reticence, there was a likelihood that surrogate endpoints would become acceptable, although some were viewed as more suitable for different diseases and cancer stages (Table 1).

The European Network of Health Technology Assessment (EUnetHTA) considers data obtained from surrogates for specific clinical endpoints if they are based on biological plausibility, supported by empirical evidence, and the associated uncertainties and limitations are explicitly explained.[38] However, to be adopted for relative effectiveness assessment purposes, specificity for disease, population, and technology would need to be shown, and the surrogate would need to be "sensible, measurable, interpretable and highly accurate in predicting the clinically relevant endpoint".[38]

The Institute for Quality and Efficiency in Health Care (IQWiG), the German HTA decision-making body, has suggested applying the surrogate threshold effect (STE) concept to estimate whether the effect on the surrogate is accompanied by an effect on the final endpoint.[12] To draw such a conclusion, the lower confidence limit of the treatment effect on the surrogate must be larger than the STE. Correlation between the surrogate and clinical endpoint is preferred but, when the relationship is unclear, they suggest using surrogate effect thresholds. For surrogate endpoints in oncology, the suggested STE threshold for high correlation (lower limit of the 95% CI) is $R \geq 0.85$, and that for low correlation (no effect; upper limit of the 95% CI) is $R \leq 0.7$.[11,12] The utility of these levels was demonstrated by Hashim et al, who used STE to assess objective response rate and PFS in relation to OS in studies involving patients with late-stage NSCLC. They concluded that OS benefit can be expected with sufficient certainty if the median objective response is ≥41.0% or PFS is longer than 4.15 months.[39]

Table 1 Key Opinion Leader Views on Suitability of Selected Surrogate Endpoints

Endpoint	Summary	Relevance	
		Adjuvant or Neoadjuvant Use	**Advanced or Metastatic Disease**
DFS	• KOLs consider DFS as to be an appropriate efficacy endpoint for adjuvant/neoadjuvant therapies in indications with good prognosis and immature survival data • Payers highlight that surrogate survival endpoints such as DFS will become more acceptable as newer products with improved efficacy are launched in earlier lines of therapy and correlation with long-term survival is established	Important relevant endpoint	Not important but relevant endpoint
MRD[a]	• KOLs expect MRD negativity to become a viable surrogate for early-line MM treatment in the absence of mature survival data, due to the correlation between MRD negativity and long-term survival. MRD is less relevant in advanced MM as it is more feasible to achieve mature PFS or OS data	Important relevant endpoint	Not important but relevant endpoint
pCR	• pCR is relevant for indications that are treated with adjuvant and neoadjuvant immunotherapy or chemotherapy before resection. pCR allows measurement of biologic activity for solid tumors and can demonstrate correlation to excellent prognosis and long-term survival • Payers agree with KOL expectations that pCR might become a key endpoint in the next 2–5 years for adjuvant/neoadjuvant therapies if correlation with long-term survival is established	Important relevant endpoint	Not important but relevant endpoint
DCR	• DCR is likely to be an appropriate supporting endpoint to PFS in metastatic indications with good prognosis and immature OS data. KOLs consider DCR to be an appropriate evolving endpoint to measure progression of metastatic disease that is not being treated with curative intent and has progressed in spite of initial therapy	Not important but relevant endpoint	Important relevant endpoint
Relapse rate	• Payers expect relapse rate to be an acceptable supporting endpoint for earlier lines of therapy but will require survival-related primary endpoints	Important relevant endpoint	Not important but relevant endpoint
Time to response	• Payers do not expect to use time to response to assess new products during HTA evaluations as type of response and duration of response are considered more important than time to first response	Not important but relevant endpoint	Not important but relevant endpoint

Notes: Based on 10 interviews with 20 stakeholders for NSCLC, prostate cancer, bladder cancer, and colorectal cancer, and three interviews with four stakeholders for multiple myeloma. [a]Relevant only to MM.

Abbreviations: DCR, disease control rate; DFS, disease-free survival; HTA, health technology assessment; KOL, key opinion leader; MRD, minimal residual disease; MM, multiple myeloma; OS, overall survival; pCR, pathologic complete response; PFS, progression-free survival.

Use of Real-World Data and Real-World Evidence

Endpoints in clinical trials are likely to differ from experiences in real-world settings. The complementary use of RWD has been proposed to broaden regulators' understanding of treatment effects in patients who would currently be excluded from trials, and that of RWE to expand appreciation of wider effects.[40,41] However, RWD are subject to several important caveats. There are recurrent issues with consistency and completeness. This is in addition to being observational in nature, derived from non-controlled environments, and not collected in relation to specific endpoints. Fears about governance and privacy are also well publicized.[42] All these factors contribute to concerns held by regulators and HTA organizations.[43] Furthermore, sources of RWD range from those established, such as medical registries, insurance databases, and electronic health records, to newer sources, such as wearable technology and social media, which can present challenges in terms of data cleaning and analysis.[42] Other factors to consider are linking between datasets (eg, primary, secondary, and community care), enrichment (the types and detail of information gathered, including, for example, QoL), costs (eg, achieving fitness for use through curation, anonymization, and simplification), and accessibility of datasets.[44] Some strides have been made in the development of frameworks to improve data quality, but further

work needs to be done to make more data intrinsically high quality, contextually appropriate, clearly represented, and accessible to potential data consumers.[44,45]

Friends of Cancer Research evaluated the performance of surrogate endpoints when analyzing RWE. In an initial pilot study, they asked six organizations to assess which endpoints for patients with advanced NSCLC who had received immune checkpoint inhibitors could be evaluated and compared across multiple data sources. They developed a framework of necessary data elements, characteristics, and definitions for real-world endpoints based on data availability in electronic health record and claims systems.[46] Friends of Cancer Research then compared consistency of US data with those extracted from the National Cancer Registration and Analysis Service (NCRAS) via the UK Cancer Analysis System for endpoints including real-world OS and time to treatment discontinuation. They found that the data for real-world time to treatment discontinuation correlated moderately to highly with real-world OS in all data sources (ρ=0.7 in CAS and ρ=0.6–0.9 in the US datasets).[47] The authors concluded "RWD can generate clinically meaningful and timely evidence on the efficacy of new cancer treatments used across diverse real-world settings", which might suggest a role in post-approval confirmatory studies. Griffith et al[48] assessed the feasibility of identifying tumor-based endpoints in RWD. In 200 patients with advanced NSCLC randomly selected from the US Flatiron database of 25,000 patients, extraction of data by clinicians supported by radiology data predicted real-world progression in 173 (87%) cases. Real-world PFS and real-world time to progression correlated well with real-world OS (ρ=0.65 and ρ=0.70, respectively). The authors concluded that data routinely collected in electronic health records could enable development and validation of surrogate endpoints of cancer progression.[48]

Guidelines on RWD and RWE use from regulators and HTA agencies are fragmented and have been created with little cross-consultation.[43] In the UK, the MHRA has recently published guidance on incorporating RWD into prospective RCTs.[49] They suggest various activities to test the data, such as feasibility studies to assess completeness of data recording and linkage in the study setting, particularly disease-specific data, and designing studies that do not alter the patients' experience of care (other than consent and randomization). NICE has also recently outlined best practices when using RWD and RWE, which suggests a role for RWD to help contextualize clinical trials by modelling the relationship between surrogate and final outcomes (including patient-reported outcomes) and to aid accurate report of post-protocol therapy.[50]

Consultations on the Use of Surrogate Endpoints

Patients' Representatives

DATA-CAN convened a group of patients' representatives to express views, for which their time was compensated. As none was required to provide personal or clinical data and the research was being conducted independently of the NHS, no informed consent for participation or review by an institutional board or ethics committee was required, in line with the guidance in the decision tool of the UK Research and Innovation, Medical Research Council, and NHS Health Research Authority.[51] The two patient groups underwent orientation sessions in which relevant terminology, such as OS and PFS, was defined in lay terms, and a glossary was provided to aid the ensuing discussions. At three meetings, patient group representatives were invited to present their thoughts, viewpoints, and experiences, and to discuss how these might contribute to the development of surrogate endpoints.

The views about OS reflected those in the 2019 national-level data from the NHS National Cancer Patient Experience Survey, in which high proportions of patients reported feeling that treatment was well explained to them (75%) and that they were involved in the decisions (81%).[52] However, respondents often felt unable to discuss worries and fears about treatment. In our consultations, the patients' representatives felt that OS was well represented and made understandable to them when discussing treatment. However, they expressed strongly that other aspects of treatment are not adequately measured and discussed in ways that would enable them to make informed choices about balancing survival against QoL outcomes that individually matter to them. They made an important distinction between overall QoL and symptom-led health-related QoL (HR-QoL). While attempts are made in clinical trials to capture HR-QoL, the common tools, such as the EuroQoL 5D and European Organization for Research and Treatment of Cancer QLQ-30 questionnaires, are not disease specific and do not fully explore or represent the short-term or long-term real-world holistic impacts on patients beyond clinical symptoms and side-effects; remaining mobile is not the same as being able to continue specific hobbies, activities, types of work, etc. Furthermore, QoL is seldom used as a primary endpoint in clinical trials,[53] missing data are

common, suitable statistical methods are not applied to test the data,[54] and data are rarely captured after study treatment ends,[55] limiting insights for long-term QoL.

Loss of any important aspect of QoL can greatly affect mental health and, in turn, disease outcomes. The patients' representatives consulted in this project strongly advocated the gathering of information on "what matters to you?" alongside "what is the matter with you?". Indeed, in a Swedish prospective cohort study of 244,261 patients followed up for 2–6 years after receiving a cancer diagnosis, 11,457 (5%) were diagnosed with mood, anxiety, and substance abuse disorders. Among these, 7236 (63%) were first-onset cases and were associated with a substantially increased risk of cancer-specific death (hazard ratio 1.82, 95% CI: 1.71–1.92).[56] Our representatives encouraged the development of appropriate surrogate endpoints to fill this gap, with support for the gathering and analysis of patient-reported outcomes such as impact on mental health, psychosocial effects, physical ability, and everyday living that could be considered along with survival.

Clinical Experts

We consulted five clinical experts in MM and 11 in LC via group meetings to obtain insights into potential differences in validation needs and standards for surrogate endpoints relevant to RWD. This group included clinicians who have worked closely with NICE and MHRA. All were contracted and compensated for providing expert advice. As for the patients' representatives, no personal or clinical data were recorded, and no informed consent for participation or review by an institutional board or ethics committee was required.

Patients with MM and LC have highly heterogeneous presentations and genetic profiles, and evaluations before treatment should consider disease-specific factors, assessment of treatment tolerance, and measurement of HR-QoL, and should involve a survival analysis. Ideal surrogate endpoints were suggested to be those that would allow early assessment and link response to the final endpoint with an established level of certainty/uncertainty and that confidence could be tested over time as more data are gathered. Potential sources of data were clinical trials and RWD.

The importance of using large datasets that can account for variation and potential confounders was emphasized. The optimum datasets would be representative across regions and include records from district general hospitals as well as larger teaching hospitals. However, a theme that unified clinical experts (and other stakeholders) was that meaningful RWD are not always available in the UK. Furthermore, data from regulatory RCTs are not always relevant to the UK health-care setting and/or are sparse. For datasets that were available, a selection was compared against a set of evaluation criteria, which are shown in Table 2.

Table 2 Assessment Criteria for Data Sources

Category	Dimension	Definition
1. Access	Information governance	Relevant if researchers other than data controllers/existing named processors use the data for secondary (research) purposes.
2. Cancer cohort(s)	MM and LC	As defined by ICD10 codes C90 (MM) and C34 (LC)
3. Support	Specialist support	Data management team available with MM and LC specialist experience
4. Data documentation	Data model, data dictionary, and provenance	Documented data model, description of data items, and description of source
5. Technical quality	Data management plan and data completeness	Description of auditing processes and data profiling results available
6. Coverage	Time period, geography, and size	Follow-up period available for each patient record, area covered, and number of patients
7. Availability	Allowable uses, frequency, timeliness, tools, and environment	Description of license agreement, including allowable users, how often the data are refreshed, time between data application and access, statistical tools available in secure analytical environment, and description of location of data for analysis
8. Value	Linkages, enrichment, and costs	Identifiers to demonstrate ability to link to other datasets, ability to link to additional bespoke datasets, and data access fees

While not exhaustive, these criteria allowed assessment of the data's potential fitness of use to validate chosen MM and LC surrogate endpoints.

Where data are sufficient, it was proposed that regression analyses would be suitable to assess potential early indicators and predictors of treatment outcomes and survivorship.

Given these restrictions, while many ideas for surrogate outcomes were discussed to enhance assessment (Tables 3–8), recommendations for immediate consideration were made for only a few where it was believed that relevant data are already recorded in the UK and/or assessments can be made in practice.

MM Clinical Endpoints

Owing to the long-term and incurable nature of MM, it was strongly recommended that surrogate endpoints would be most informative where assessment of OS is most challenging (first-line or second-line therapy) or most likely to be confounded by other factors (eg, multiple subsequent lines of therapy). Time to next treatment ([TTNT] as a marker of disease progression), relapse kinetics measured by rate of rise in paraprotein or serum-free light chain (sFLC) concentrations, durability of treatment response, and HR-QoL and QoL were judged to be most relevant (Box 1).

TTNT seems to be well reported in registries, which has allowed its inclusion in RWD assessments of MM treatments.[57–65] It directly reflects the time during which patients do not require a subsequent line of therapy, which can be an important measure of how well a treatment is controlling the disease and, to some extent, how well it is tolerated. The clinical experts recommended that TTNT data should be routinely collected in clinical trials for planned subgroup analyses to evaluate the effects of key confounders of this endpoint (eg, patient demographics, disease-specific factors, and physician-related factors). Additionally, as many of the tests and methods used in RCTs will not be feasible for routine clinical care, the clinical experts suggested that surrogate endpoints that are readily available in RWD be measured during clinical trials, which would ensure that the trial evidence is comparable with RWD. Of note, though when reviewing these discussions, the patients' representatives highlighted that it should be explored whether factors, such as a patient choosing to delay intensive treatments (eg, high-dose chemotherapy followed by autologous stem cell transplantation, which can be painful and debilitating), affect the validity of TTNT for MM.

While clinical trials remain relevant in all groups, they are particularly appropriate for relapsed patients, who form an extremely heterogeneous subgroup. OS should be the preferred endpoint in late-stage trials owing to short life expectancy, limited confounding, and an early available comparison of potential OS benefit. However, the clinical experts recommended including explorative subgroup analyses of biochemical markers of relapse (eg, rate of rise in paraprotein or sFLC), which will be of academic interest. RWD collection and recording are key potential areas for improvement.

The use of composite clinical trial endpoints that include biochemical indicators of the functional consequences of disease progression or organ damage was also proposed. These could include, for example, renal function, hemoglobin levels, calcium levels, and skeletal-related events. However, any composite endpoints would require extensive validation in controlled clinical trial environments before any wider adoption, and they would be highly unlikely to replace OS as a primary endpoint.

LC Clinical Endpoints

The clinical expert group for LC felt that surrogate endpoints would be particularly appropriate for the assessment of early-stage disease where OS is least meaningful due to potential length of survival and potential for curative intention of treatment. DFS was recommended for stage I and II disease to indicate the end of a disease-free period (eg, due to recurrence, need for a new therapy, or death) and could be collected using codes in NCRAS and the Systemic Anti-Cancer Therapy Dataset that have already been captured for cancer registration in NSCLC. Fiteni et al[66] found that the available evidence supported DFS as a surrogate marker for OS when considered in the IQWiG criteria framework for surrogate endpoints. Among 20 studies, in trials of adjuvant chemotherapy, the correlation between DFS and OS was 0.83 at the individual level (95% CI 0.83–0.83) and 0.92 at trial level (95% CI 0.88–0.95).[66] As for MM, more robust evaluation of HR-QoL was recommended, with broadening of data to capture overall QoL being added when suitable tools are developed (Box 2).

Table 3 Clinical Expert Discussion of Potential Surrogate Endpoints in MM: Transplant-Eligible Patients

Key Measure	Clinical Trial Endpoint	RWD and RWE Discussions	Currently Suitable as a Surrogate Endpoint?
Survival	OS	• Improved survival means that trials in this cohort risk will be obsolete by the time of survival data readout. • Subsequent lines of treatment are poorly reported in clinical trials. • OS is not meaningful due to the potential length of survival and the impact of subsequent lines of treatment. • RWD should be improved by measurement of disease/morbidity-related survival and possibly quality-adjusted and age-related relative survival. • Patients' perspectives on balancing HR-QoL with survival may differ.	No, but RW assessment of OS should be improved
Duration of response	PFS according to IMWG categories	• PFS in first-line management offers greatest survival benefit, particularly the duration of first response versus early relapse. • Numerical improvement in PFS is not relevant in an RW setting as it does not directly impact clinical care. • Quantitative paraprotein/sFLC assessments that are routinely collected must be easier to obtain.	Yes
Depth of response/ kinetics of relapse	Response rate according to the International Uniform Response Criteria for Multiple Myeloma, including sCR, CR, VGPR, and PR.	• Studies suggest that depth of response is less important than sustainability of VGPR. • Biochemical laboratory RWD on relapse kinetics could be important, but easier access is required. • Time without treatment is important to patients.	Yes
Residual tumor burden	MRD according to IMWG criteria.	• MRD is currently measured at only one time point as standard. • PET-CT MRD is expensive and not consistently available in the UK. • Measurement of increasing depth of response or capture of relapse earlier would be ideal but change of clinical practice is not currently likely. • The current frequency of assessments for MRD is unlikely to be applicable to the clinical setting due to patient and clinical burden and current need for invasive bone-marrow biopsies.	No
Negative impact of therapy	Safety assessments (ADRs) and treatment tolerance (dose reductions/ treatment discontinuation)	• Includes renal impairment, neuropathy, immunodeficiency, infection, thrombosis, and hospitalizations. • Assess ToT in first-line assessments to evaluate the impact of toxicities and consider timings of cycles, treatment delays, and dose reductions. • Trial composite endpoints could include relevant assessments. • Potential to also use recovery rates from HDT-ASCT treatment, including markers of frailty, but these assessments are not yet validated	Yes, including as part of a composite/multicomponent endpoint
HR-QoL (including impact of treatment and disease progression)	EQ-5D/EORTC-QLQ30 and PROs	• Desired by patients, but EQ-5D not regularly used in clinical care. • Other options might include wearable technologies (eg, fitness/mobility measures) and clinical measures (eg, hospitalization, length of stay, infection, and health-resource utilization). • Access to GP records and natural language processing of clinical records could improve data capture.	Yes, for all patients, but lack of longitudinal HR-QoL data is a key issue

Abbreviations: ADR, adverse drug reaction; CR, complete response; EORTC, European Organization for Research and Treatment of Cancer; EQ-5D, EuroQoL 5D tool; HR-QoL, health-related quality of life; HDT-ASCT, high-dose chemotherapy followed by autologous stem cell transplantation; GP, general practitioner; IMWG, International Myeloma Working Group; MRD, minimal residual disease; OS, overall survival; PET-CT, positron emission tomography-computed tomography; PFS, progression-free survival; PR, partial response; PRO, patient-reported outcome. RW, real-world; RWD, real-world data; RWE, real-world evidence; sCR, stringent complete response; sFLC, serum free light chain; ToT, time on treatment; VGPR, very good partial response.

Table 4 Clinical Expert Discussion of Potential Surrogate Endpoints in Multiple Myeloma: Transplant-Ineligible Patients

Key Measure	Clinical Trial Endpoint	RWD and RWE Discussions	Currently Suitable as a Surrogate Endpoint?
Survival	OS	• OS is often shortened by increased age and/or frailty. • Comorbidities are likely, so cause of death is important to establish. • Assessment of disease/morbidity-related survival is essential. • Also consider quality-adjusted and age-related survival.	No, but RW assessment of OS should be improved
Duration of response	PFS according to IMWG categories	• Less-pronounced correlation of OS in these patients is due to increased confounding factors. Appropriate adjustments (eg, age and/or frailty score) could help. • Co-morbidities are also increasingly important. • TTNT may be an option to explore. • Composite trial endpoints could incorporate toxicities and treatment discontinuation. • Consider including supportive therapy, such as transfusion, medications/analgesia, immunoglobulins, and antibiotics. • Timing of referral to palliative care could be informative.	Yes
Depth of response/kinetics of relapse	Response rate according to International Uniform Response Criteria for Multiple Myeloma, including sCR, CR, VGPR, and PR.	• Biochemical relapse kinetics could be important but easier access is required.	Yes
Residual tumor burden	MRD according to IMWG criteria	• MRD is currently measured at only one time point as standard. • PET-CT MRD is expensive and not consistently available in the UK. • Measurement of increasing depth of response or capture of relapse earlier would be ideal but change of clinical practice is not currently likely. • The current frequency of assessments for MRD is unlikely to be applicable to the clinical setting due to patient and clinical burden and current need for invasive bone-marrow biopsies.	No
Negative impact of therapy	Safety assessments (ADRs) and treatment tolerance (dose reductions/ treatment discontinuation)	• Includes renal impairment, neuropathy, immunodeficiency, infection, thrombosis, and hospitalizations. • Assess ToT in first-line assessments to evaluate impact of toxicities and consider timings of cycles, treatment delays, and dose reductions. • Trial composite endpoints could include relevant assessments. • Potential to also use recovery rates from HDT-ASCT treatment, including markers of frailty, but these assessments are not yet validated	Yes, including as part of a composite/ multicomponent tool
HR-QoL (including impact of treatment and disease progression)	EQ-5D/EORTC-QLQ30	• Desired by patients, but EQ-5D not regularly used in clinical care. • Other options might include wearable technologies (eg, fitness/mobility measures), clinical measures (eg, hospitalization, length of stay, infection, and health-resource utilization). • Access to GP records and natural language processing of clinical records could improve data capture.	Yes, for all patients

Abbreviations: ADR, adverse drug reaction; CR, complete response; EORTC, European Organization for Research and Treatment of Cancer; EQ-5D, EuroQoL 5D tool; GP, general practitioner; HDT-ASCT, High-dose chemotherapy followed by autologous stem-cell transplantation; HR-QoL, health-related quality of life; IMWG, International Myeloma Working Group; MRD, minimal residual disease; OS, overall survival; PET-CT, positron emission tomography-computed tomography; PFS, progression-free survival; PR, partial response; RW, real-world; RWD, real-world data; RWE, real-world evidence; sCR, stringent complete response; ToT, time on treatment; TTNT, time to next treatment; VGPR, very good partial response.

https://doi.org/10.2147/CMAR.S441359

Table 5 Clinical Expert Discussion of Potential Surrogate Endpoints in Multiple Myeloma: Relapsing or Refractory Patients

Key Measure	Clinical Trial Endpoint	RWD and RWE Discussions	Currently Suitable as a Surrogate Endpoint?
Survival	OS	• OS is usually shorter in this population and more likely to be the most meaningful outcome measure, if readout of survivorship is optimized. The justification for other endpoints, particularly in relapsing/refractory patients, is less compelling.	No, but RW assessment of OS should be improved
Duration of response	PFS2	• Moderate to strong correlation with OS is possible but not proven. • Potentially too many confounders in this population to be meaningful, but the relapsing patients should still be included in a TTNT assessment.	Yes, but only for relapsing patients
Kinetics of relapse	Response rate according to International Uniform Response Criteria for Multiple Myeloma, including sCR, CR, VGPR, and PR.	• Biochemical laboratory RWD on relapse kinetics could be important but easier access is required.	Yes, but only for relapsing patients
Negative impact of therapy	Safety assessments (ADRs) and treatment tolerance (dose reductions/ treatment discontinuation)	• Includes renal impairment, neuropathy, immunodeficiency, infection, thrombosis, and hospitalizations. • Assess ToT in first-line assessments to evaluate the impact of toxicities and consider timings of cycles, treatment delays, and dose reductions. • Trial composite endpoints could include relevant assessments. • Potential to also use recovery rates from HDT-ASCT treatment, including markers of frailty, but these assessments are not yet validated.	Yes, including as part of a composite/ multicomponent tool
HR-QoL (including impact of treatment and disease progression)	EQ-5D/EORTC-QLQ30	• Desired by patients, but EQ-5D not regularly used in clinical care. • Other options might include wearable technologies (eg, fitness/mobility measures), clinical measures (eg, hospitalization, length of stay, infection, and health-resource utilization). • Access to GP records and natural language processing of clinical records could improve data capture.	Yes, for all patients

Abbreviations: ADR, adverse drug reaction; CR, complete response; EORTC, European Organization for Research and Treatment of Cancer; EQ-5D, EuroQoL 5D tool; GP, general practitioner; HR-QoL, health-related quality of life; OS, overall survival; PFS2, time from randomization to progression on second-line therapy; PR, partial response; RW, real-world; RWD, real-world data; RWE, real-world evidence; sCR, stringent complete response; TTNT, time to next treatment; ToT, time on treatment; VGPR, very good partial response.

Potential challenges to the use of DFS as a surrogate endpoint are the lack of access to molecular marker RWD and imaging data potentially being poor dependent on image scheduling and recording in NCRAS. The individual and heterogeneous nature of LC is becoming increasingly recognized, and these data will be essential to effectively stratify and identify specific patient cohorts based on the genetic drivers of their disease. In this vein, stage III disease was considered by the clinical experts to be too heterogeneous and complex group to control for all the potential confounding factors and, therefore, the use of these surrogate endpoints was not extended to this subgroup.

Table 6 Discussion of Additional Potential Surrogate Endpoints in MM Based on RWD and RWE

Key Measure	RWD and RWE Discussion	Currently Suitable as a Surrogate Endpoint?
Markers of immunological fitness	• Considered important, potentially as a component of other confounding factors.	Maybe, as part of assessment of previous treatment
Quantitative assessments of physical fitness	• Important HR-QoL parameter that could be considered (eg, what are you able to do whilst on treatment?).	Maybe, as part of assessment of HR-QoL
Biochemical identification of frail patients (eg, MRP)	• Risk-adapted disease management is becoming increasingly important. • MRP could offer a biochemical indicator but has not been validated as an outcome measure.	No
Functional consequences of disease progression (renal/anemia)	• May be a meaningful assessment as part of a composite assessment in combination with relapse rate. • Biochemical assessment of organ damage was proposed, including renal function, hemoglobin, calcium, and bone disease (eg, skeletal-related events)	Yes, as part of biochemical assessment of relapse kinetics and disease progression (as a composite or multicomponent endpoint)

Abbreviations: HR-QoL, health-related quality of life; MRP, UK Myeloma Research Alliance Risk Profile. RWD, real-world data; RWE, real-world evidence.

Minimal Residual Disease

MRD has received considerable attention as a surrogate endpoint in MM[67,68] and LC (assessed by measurement of circulating tumor DNA).[69–71] However, measurement is reliant on emerging technologies in cytometry, PCR, and next-generation sequencing that are time-consuming, expensive, and/or not readily available. How MRD status would contribute to clinical practice and which method is preferable have also yet to be clearly established in these diseases.[72] MRD status might be of value in detecting relapse after surgery in LC,[73] and has been demonstrated to have prognostic value in both MM and LC.[74,75] Yet, this potential surrogate endpoint will continue to be somewhat limited for MM until benefits to the patients in terms of influence on choice or timing of therapy are demonstrated. RWD for this surrogate endpoint were deemed immature, largely owing to inconsistent timing of assays and variations in the techniques used. However, although MRD was not recommended for either disease currently, it was reserved for special mention as a potential option in these and other cancers in the future as the technology improves. Indeed, MRD-directed therapeutic decision-making is already a clinical reality in acute and chronic myeloid leukemia. In acute myeloid leukemia, MRD negativity is associated with improved outcomes[76] and is considered when making decisions about the intensity of treatment,[77] treatment intensification,[78] and the need for stem-cell transplantation. In chronic myeloid leukemia, MRD monitoring through quantitative PCR for *BCR-ABL1* transcripts is a cornerstone of treatment management, informing decisions on treatment intensity, changes, and, in some cases, the safe discontinuation of therapy for patients who achieve deep and sustained molecular responses.[79]

Industry

Industry consultations indicated that surrogate endpoints are viewed as being beneficial, providing opportunities for clearer understanding of disease characteristics and, thereby, predictive or prognostic factors, outcomes (including toxicity), and options for treatment (duration, sequences, and management of adverse reactions). Schievink et al reported similar enthusiasm by industry stakeholders.[80]

Importance was placed on the opportunities RWD provide to obtain information on outcomes in the many cancer patients not eligible for clinical trials. Other noted benefits were the ability to assess longer-term outcomes in disease-free patients, evaluate real-world mortality compared with the general population (or to identify a standardized mortality ratio), and explore potential correlations between RWE and clinical trial surrogate outcomes, which could accelerate understanding and acceptance of early endpoints.

https://doi.org/10.2147/CMAR.S441359 DovePress

Table 7 Clinical Expert Discussion of Potential Surrogate Endpoints in Lung Cancer: Stage I and II Disease

Key Measure	Clinical Trial Endpoint	RWD and RWE Discussions	Currently Suitable as a Surrogate Endpoint?
Survival	OS	• OS is least meaningful in early-stage disease due to curative intent of treatment and potentially long survival. It is also confounded by multiple factors in RWD, many of which cannot be effectively measured. • Measurement of disease/morbidity-related survival is essential. • Recurrence-free survival can be very difficult to determine from RWD on mortality. • Quality-adjusted survival and age-related relative survival should also be considered. • Mortality rates and mortality per person-years (age standardized) could be informative. • Patients want to know the likelihood of survival by time and/or the probability of death or disease status.	No, but RW assessment of OS should be improved
Recurrence of disease/ duration of response	DFS (increasingly accepted by regulators)	• Important to determine effective personalized adjuvant therapies; must be contextualized with data on patients' characteristics and comorbidities. • Frequency of scans to detect disease recurrence needs to be improved. • Need to identify appropriate correlation factor to OS (eg, 2 years or 5 years).	Yes, so long as it shows any magnitude of change/ differentiation
Pathological assessments	Pathological stage/ prognosis	• TNM-PS staging remains the standard. • Combined endpoints could be informative for assessing predictive potential of key assessments (eg, vascular invasion, volume doubling) but are still not clearly defined. • Reporting and standardization of data recording for other factors could be improved.	No
Residual disease assessments	MRD negativity based on assessments of ctDNA	• Interesting preliminary data on the utility of ctDNA assessments but is still in very early stages of validation and potential clinical adoption. • ctDNA has the clear potential to be much more sensitive at detecting the presence of disease than a scan.	No, but should undergo further exploration
HR-QoL	EQ-5D/EORTC-QLQ30	• The link between DFS and long-term HR-QoL is important for early-stage adjuvant treatments. Could consider correlating DFS/PFS to HR-QoL, particularly in early disease. • PROs are not captured sufficiently to understand true impact on HR-QoL. • Classification by "health states" and assessment in studies might be useful as a surrogate endpoint for HR-QoL.	Yes, for all patients

Other prognostic factors	Confounders often "selected out" by trial inclusion and exclusion criteria	• It may be difficult to subcategorize prognostic factors beyond TNM-PS and even in composite measures this indicator appears to outweigh other factors. • Early detection of tumors is a known prognostic factor and potentially indicates a more indolent tumor with better survival outcomes. Achieving earlier diagnosis is important (eg, through more effective screening) and tumor aggressiveness should be included in the prognostic assessment. • Type of surgery/radiotherapy are also known as prognostic factors for survival. • PS is favored over frailty assessments. • Collection methods need to be simple but ensure that the nuances may be captured.	It was suggested that baseline characteristics and other key indicators and indolence/ aggressiveness of the tumor could be used to place surrogate endpoints such as DFS in context

Abbreviations: ctDNA, circulating tumor DNA; DFS, disease-free survival; EORTC, European Organization for Research and Treatment of Cancer; EQ-5D, EuroQoL 5D tool; HR-QoL, health-related quality of life; MRD, minimal residual disease; OS, overall survival; PRO, patient-reported outcome; PS, performance status; RW, real-world; RWD, real-world data; RWE, real-world evidence; TNM-PS, tumor node metastasis status and performance status.

https://doi.org/10.2147/CMAR.S441359

Table 8 Clinical Expert Discussion of Potential Surrogate Endpoints in Lung Cancer: Stage III Disease

Key Measure	Clinical Trial Endpoint	RWD and RWE Discussions	Currently Suitable as a Surrogate Endpoint?
Survival	OS	• OS is least meaningful as curative intent remains in this population. It is also confounded by multiple factors in RWD, many of which cannot be effectively measured.	Not a surrogate endpoint, but RW assessment of OS should be improved
Recurrence of disease	DFS	• Important to determine effective personalized adjuvant therapies; must be contextualized with data on patients' characteristics and comorbidities, but morphological changes due to the treatments can make the imaging difficult to interpret. • Frequency of scans to detect disease recurrence needs to be improved. • Need to identify appropriate correlation factor to OS (eg, 2 years or 5 years). • Understanding of wider imaging modalities needs to be improved (eg, PET standardized uptake values). • Improved data on proportion of patients responding to treatment/extent of response would be beneficial.	No
Pathological response	Pathological response	• Whether major versus complete pathological response is most appropriate is unclear, and assessment approaches differ. • Pathological response assessments take a long and require specific training, which affects applicability in clinical practice. What value data collected in a clinical trial setting and used in the regulatory approval process offer to the clinical setting is unclear. • Pathological assessment measures response only in resected tumors but not residual tumor or micrometastatic disease	No
Residual disease assessments	MRD negativity based on assessments of ctDNA	• Interesting preliminary data on the utility of ctDNA assessments but is still in very early stages of validation and potential clinical adoption. • ctDNA has the clear potential to be much more sensitive at detecting the presence of disease than a scan.	Should undergo further exploration
Negative impact of therapy	Surgical/ radiotherapy complications, ADRs, and treatment tolerance	• Composite RW surrogates that offer a more comprehensive evaluation of the benefit and adverse outcomes of treatments could provide clinical value but will be challenging to undertake.	Likely too complicated to take forward

Overall QoL (including impact of treatment and disease progression)	EQ-5D/EORTC-QLQ30	• Lack of longitudinal data is a key issue. If measured in trials, data are generally not reported or used. • QALYs might offer a direct assessment pathway but require longer-term measurements of HR-QoL. • Collection methods need to be simple but ensure that the nuances may be captured. • How regulators value HR-QoL needs clarification. • New technologies, such as smart phone applications and wearables, could be important for capture of QoL data. • Access to GP records may be useful and natural language processing of clinical records could be considered to improve capture of data.	Yes, for all patients

Abbreviations: ADR, adverse drug reaction; ctDNA, circulating tumor DNA; DFS, disease-free survival; EORTC, European Organization for Research and Treatment of Cancer; EQ-5D, EuroQoL 5D tool; GP, general practitioner; HR-QoL, health-related quality of life; QoL, quality of life; MRD, minimal residual disease; OS, overall survival; PET, positron emission tomography; PS, performance status; QALY, quality-adjusted life-years. RW, real-world; RWD, real-world data; RWE, real-world evidence; TNM-PS, tumor node metastasis status and performance status.

https://doi.org/10.2147/CMAR.S441359

Box 1 Surrogate Endpoint Recommendations for Multiple Myeloma

The following surrogate endpoints were suggested by the clinical experts for further assessment:
In patients receiving first-line treatment (transplant eligible and ineligible) or experiencing an initial relapse from therapy:
1. Time to next treatment as an indicator of disease progression
2. Biochemical relapse kinetics by rate of rise of paraprotein or serum-free light chain, potentially alongside incorporation of biochemical indicators of disease progression/organ damage as a composite assessment
In relapsed patients:
3. Additional measures to evaluate how well the patient tolerated/responded to previous treatment, eg, as part of a composite endpoint
In all patients:
4. Improved collection, evaluation, and use of longitudinal health-related quality of life (in the absence of overall quality-of-life surrogate endpoints) at all stages of treatment. In the absence of patient-reported health-related quality of life, potential proxies for real-world data could be hospitalization, infection, and health-resource utilization.

Box 2 Surrogate Endpoint Recommendations for Lung Cancer

In early-stage disease (stage I and II):
1. Disease-free survival through either the capture of i) a documented recurrence (scan or positive pathology); ii) a measure of the disease-free period (eg, treatment change or initiation of a new therapy), or iii) death. (May also include assessment of patient notes in electronic health records.)
In all patients:
2. Improved collection, evaluation, and use of longitudinal health-related quality of life data (in the absence of overall quality-of-life surrogate endpoints) for all patients, but particularly during end-stage disease. In the absence of patient-reported health-related quality of life, potential proxies for real-world data could be health needs assessments, number of drug therapies, and frailty indices.

The use of RWD was generally supported to provide comparative data in single-arm trials or for diseases with small numbers of patients, support accelerated approval and outcomes-based pricing by enabling post-marketing analysis and contribute to RCT design.

Several challenges to using surrogate endpoints were mentioned. Observational data cannot indicate causal associations and have high potential for bias, which could create challenges to revealing correlations between populations of patients. RWD can be affected by structure, content, and coding systems, fragmentation, completeness, and timeliness. Data accessibility due to governance and payment structures was also cited as an issue.

Access to imaging and genomic information with sufficient quantity and quality to enable the use of artificial intelligence (eg, supervised or semi-supervised machine learning) was thought to be an important issue due to the high monetary and time costs involved in generating new sets. In line with the other groups consulted, industry representatives underscored that QoL metrics are poorly recorded. However, again, they noted that the efforts to collect additional datasets could lead to additional stress on health systems.

Finally, although it was not suggested that surrogate endpoints in RWD could completely replace clinical trial data, it was felt that they could help to overcome some of the challenges of continued requests for prospective trial data by regulators and HTA appraisers. As mentioned by the clinical experts, a suggestion to improve the reliability of RWD – and perhaps to strengthen the power of trials – was to incorporate relevant surrogate endpoints, such as biomarkers and tests, into composite primary endpoints.

Conclusions

The discussions related to this work suggest that RWD and RWE, at least in theory, provide potential opportunities for further evaluation of surrogate endpoints. Already, findings derived from the use of well-populated large datasets in the USA have provided examples where specific real-world surrogate endpoints, such as PFS and time to progression, have shown good correlation with real-world OS.[47,48] In the UK, challenges related to gaps in documentation, preservation, and accessibility of critical data are important to address, as is the issue of identifying the types of data that various

stakeholders find most valuable. Standardization of terminology, frameworks for recording of data, and the development of tools to capture broader characteristics, such as QoL, need to be further explored, validated, and implemented. However, although it will take time to create sufficient new datasets, enhance those existing, and to link the data fully, we believe that this is an opportune moment to explore ways to improve completeness, quality, and clinical and academic relevance of RWD and maximize their relevance to patients. In turn, appropriately designed studies may facilitate the development of biologically plausible and patient-focused surrogate endpoints that could support the requirements of all stakeholders. Additionally, the patients' representatives we engaged with offered clear perspectives on how RWD could substantially enhance the treatment experience in ways not adequately captured by current clinical trials. Consequently, establishing a data group consisting of patients and caregivers would markedly improve both the availability and the relevance of pertinent data.

Acknowledgments

We thank Lucinda Billingham, Ceri Bygrave, David Cairns, John Conibear, Alistair Greystoke, Allan Hackshaw, Keith Kerr, Richard Lee, Sarah Lawless, Dionysis Papadatos-Pastos, Helen Powell, Neil Rabin, and Paresh Sewpaul for participation in workshops to discuss surrogate endpoints in cancer. The feasibility study discussed in this article was co-funded by Janssen-Cilag Ltd and The Association of the British Pharmaceutical Industry. GC and JC were supported by the National Institute for Health Research (NIHR) infrastructure at Leeds, UK. NN is supported by a Medical Research Council Clinical Academic Research Partnership (MR/T02481X/1). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. We thank Rachel Ashton for writing support, which was funded by the study sponsor, and Paresh Sewpaul, Janssen-Cilag Ltd, for support across the study.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Disclosure

Dr Khalid Siddiqui is a current employee of Johnson and Johnson Innovative Medicine, UK. Dr David Baldwin reports personal fees, non-financial support from Janssen during the conduct of the study; personal fees from AstraZeneca, personal fees from MSD, and personal fees from Roche; and honoraria from Bristol Myers Squibb, Roche, Astra Zeneca, and MSD outside the submitted work. Dr Jonathan Carmichael reports personal fees from Janssen, during the conduct of the study. Dr Gordon Cook reports grants, personal fees from Janssen, grants, personal fees from Takeda, grants, personal fees from BMS, personal fees from Pfizer, and personal fees from Amgen, during the conduct of the study. Dr Neal Navani reports grants from Medical Research Council, personal fees, non-financial support from Amgen, grants, personal fees, non-financial support from Astra Zeneca, personal fees, non-financial support from Boehringer Ingelheim, personal fees from Bristol Meyers Squibb, personal fees from EQRx, personal fees, non-financial support from Fujifilm, personal fees from Guardant Health, personal fees from Intuitive, personal fees from Janssen, grants, personal fees from Lilly, personal fees from Merck Sharp Dohme, personal fees, non-financial support from Olympus, personal fees from Roche, during the conduct of the study. Mr James Peach reports grants from Jansen & Jansen, grants from ABPI, during the conduct of the study; Shareholding from Univ8 Genomics, personal fees from IQVIA, personal fees from Roche, personal fees from ABPI, personal fees from Health Data Research UK, personal fees from Sensyne Health, personal fees from Novartis, personal fees from Medicines Discovery Catapult, outside the submitted work; and Voluntary (unpaid) advisor to Cancer Research UK and sit on funding committee. Dr Nicola Allen-Delingpole reports that ABPI is a trade association representing the pharmaceutical industry in the UK funded by annual subscriptions from member companies from ABPI, outside the submitted work; Dr Cicely Kerr was an employee of Janssen-Cilag at the time the work was commissioned by this company and currently holds shares in the company. The authors report no other conflicts of interest in this work.

References

1. Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study. *Lancet*. 2015;385(9974):1206–1218. doi:10.1016/S0140-6736(14)61396-9
2. Delgado A, Guddati AK. Clinical endpoints in oncology - a primer. *Am J Cancer Res*. 2021;11(4):1121–1131.
3. Boshuizen J, Peeper DS. Rational Cancer Treatment Combinations: an Urgent Clinical Need. *Mol Cell*. 2020;78(6):1002–1018. doi:10.1016/j.molcel.2020.05.031
4. He J, Morales DR, Guthrie B. Exclusion rates in randomized controlled trials of treatments for physical conditions: a systematic review. *Trials*. 2020;21(1):228. doi:10.1186/s13063-020-4139-0
5. Capobianco E. High-dimensional role of AI and machine learning in cancer research. *Br J Cancer*. 2022;126(4):523–532. doi:10.1038/s41416-021-01689-z
6. Strzebonska K, Waligora M. Umbrella and basket trials in oncology: ethical challenges. *BMC Med Ethics*. 2019;20(1):58. doi:10.1186/s12910-019-0395-5
7. Del Paggio JC, Berry JS, Hopman WM, et al. Evolution of the Randomized Clinical Trial in the Era of Precision Oncology. *JAMA Oncol*. 2021;7(5):728–734. doi:10.1001/jamaoncol.2021.0379
8. Shafrin J, Brookmeyer R, Peneva D, et al. The value of surrogate endpoints for predicting real-world survival across five cancer types. *Curr Med Res Opin*. 2016;32(4):731–739. doi:10.1185/03007995.2016.1140027
9. Etekal T, Koehn K, Sborov DW, et al. Time-to-event surrogate end-points in multiple myeloma randomised trials from 2005 to 2019: a surrogacy analysis. *Br J Haematol*. 2023;200(5):587–594. doi:10.1111/bjh.18568
10. Dawoud D, Naci H, Ciani O, Bujkiewicz S. Raising the bar for using surrogate endpoints in drug regulation and health technology assessment. *BMJ*. 2021;374:n2191. doi:10.1136/bmj.n2191
11. 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;41(4):439–452. doi:10.1177/0272989X21994553
12. Institute for Quality and Efficiency in Health Care. IQWiG Reports – commission No. A10-05. Validity of surrogate endpoints in oncology; 2011. Available from: https://www.iqwig.de/download/a10-05_executive_summary_v1-1_surrogate_endpoints_in_oncology.pdf. Accessed May 3, 2022.
13. Downing NS, Aminawung JA, Shah ND, Krumholz HM, Ross JS. Clinical trial evidence supporting FDA approval of novel therapeutic agents, 2005–2012. *JAMA*. 2014;311(4):368–377. doi:10.1001/jama.2013.282034
14. Bruce CS, Brhlikova P, Heath J, McGettigan P. 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 Med*. 2019;16(9):e1002873. doi:10.1371/journal.pmed.1002873
15. Committee for Medicinal Products for Human Use (CHMP). Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies; 2016. Available from: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf. Accessed May 3, 2022.
16. FDA. PDUFA reauthorization performance goals and procedures fiscal years 2018 through 2022; 2022. Available from: https://www.fda.gov/media/99140/download. Accessed May 3, 2023.
17. Medicines and Healthcare products Regulatory Agency. Consultation document: MHRA draft guidance on randomised controlled trials generating real-world evidence to support regulatory decisions. GOV.UK; 2021. Available from: https://www.gov.uk/government/consultations/mhra-draft-guidance-on-randomised-controlled-trials-generating-real-world-evidence-to-support-regulatory-decisions/consultation-document-mhra-draft-guidance-on-randomised-controlled-trials-generating-real-world-evidence-to-support-regulatory-decisions. Accessed May 3, 2022.
18. Chen EY, Joshi SK, Tran A, Prasad V. Estimation of study time reduction using surrogate end points rather than overall survival in oncology clinical trials. *JAMA Intern Med*. 2019;179(5):642–647. doi:10.1001/jamainternmed.2018.8351
19. Kleijnen S, Lipska I, Alves TL, et al. Relative effectiveness assessments of oncology medicines for pricing and reimbursement decisions in European countries. *Ann Oncol*. 2016;27(9):1768–1775. doi:10.1093/annonc/mdw233
20. FDA. Table of Surrogate Endpoints That Were the Basis of Drug Approval or Licensure; 2022. Available from: https://www.fda.gov/drugs/development-resources/table-surrogate-endpoints-were-basis-drug-approval-or-licensure. Accessed May 3, 2022.
21. Satherley A, Chang E, Awasthy S, Michel S. Understanding payer sensitivities when considering the use of surrogate endpoints to substantiate clinical value propositions: country differences between England, Germany, and the U.S. In: *The Evidence Forum*. Spring.Vol. 2017; 2017. Available from: https://www.evidera.com/wp-content/uploads/2017/05/Surrogate-Endpoints-Used-in-Health-Technology-Assessments.pdf. Accessed June 27, 2024.
22. Kordecka A, Walkiewicz-żarek E, Łapa J, Sadowska E, Kordecki M. Selection of endpoints in clinical trials: trends in European marketing authorization practice in oncological indications. *Value Health*. 2019;22(8):884–890. doi:10.1016/j.jval.2019.03.007
23. Medicines and healthcare products regulatory agency. Guidance: innovative licensing and access pathway. GOV.UK; 2021. Available from: https://www.gov.uk/guidance/innovative-licensing-and-access-pathway. Accessed May 3, 2022.
24. Scottish Medicines Consortium. March 2022 decisions news release. Scottish Medicines Consortium; 2022. Available from: https://www.scottishmedicines.org.uk/about-us/latest-update/march-2022-decisions-news-release/. Accessed March 20, 2024.
25. Grigore B, Ciani O, Dams F, et al. Surrogate endpoints in health technology assessment: an international review of methodological guidelines. *Pharmacoeconomics*. 2020;38(10):1055–1070. doi:10.1007/s40273-020-00935-1
26. 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*. 2017;20(3):487–495. doi:10.1016/j.jval.2016.10.011
27. NICE. Brentuximab vedotin for treating CD30-positive Hodgkin lymphoma; 2018. Available from: https://www.nice.org.uk/guidance/ta524/chapter/1-Recommendations. Accessed February 20, 2024.
28. Cortazar P, Zhang L, Untch M, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. *Lancet*. 2014;384(9938):164–172. doi:10.1016/S0140-6736(13)62422-8
29. Pan-Canadian oncology drug review. Pertuzumab (Perjeta) NBC – final recommendation; 2015. Available from: https://www.cadth.ca/sites/default/files/pcodr/pcodr_pertuzumab_perjeta_nbc_fn_rec.pdf. Accessed February 15, 2024.
30. Rizzello E, Leopaldi C, Massó B, et al. Acceptance of progression-free survival 2 (PFS2) by EU5 and Canadian health authorities for cancer drug reimbursement. *Value Health*. 2022;25(1 (suppl)):S184(abstrPOSC258). doi:10.1016/j.jval.2021.11.896
31. NICE. NICE health technology evaluations: the manual. Process and methods [PMG36]. *Natl Inst Health Care Excell*. 2022;2022:181.

32. NICE. Osimertinib for adjuvant treatment of EGFR mutation-positive non-small-cell lung cancer after complete tumour resection. Technology appraisal guidance [TA761]; 2022. Available from: https://www.nice.org.uk/guidance/ta761/chapter/3-Committee-discussion. Accessed June 2, 2022.
33. NICE. Daratumumab in combination for untreated multiple myeloma when a stem cell transplant is suitable. Technology appraisal guidance [TA763]; 2022. Available from: https://www.nice.org.uk/guidance/ta763/chapter/3-Committee-discussion. Accessed June 2, 2022.
34. Bouvy J How 'surrogate outcomes' influence long-term health outcomes | blogs | news. NICE; 2023. Available from: https://www.nice.org.uk/news/blog/how-surrogate-outcomes-influence-long-term-health-outcomes. Accessed February 20, 2024.
35. Cooper K, Tappenden P, Cantrell A, Ennis K. A systematic review of meta-analyses assessing the validity of tumour response endpoints as surrogates for progression-free or overall survival in cancer. *Br J Cancer*. 2020;123(11):1686–1696. doi:10.1038/s41416-020-01050-w
36. Pasalic D, McGinnis GJ, Fuller CD, et al. Progression-free survival is a suboptimal predictor for overall survival among metastatic solid tumor clinical trials. *Eur J Cancer Oxf Engl 1990*. 2020;136:176–185. doi:10.1016/j.ejca.2020.06.015
37. IQVIA Institute. Evolving oncology endpoints; 2021. Available from: https://www.iqvia.com/insights/the-iqvia-institute/reports/evolving-oncology-endpoints. Accessed May 3, 2022.
38. EUnetHTA. Endpoints used in relative effectiveness assessment: surrogate endpoints; 2015. Available from: https://www.eunethta.eu/wp-content/uploads/2018/01/Endpoints-used-in-Relative-Effectiveness-Assessment-Surrogate-Endpoints_Amended-JA1-Guideline_Final-Nov-2015.pdf. Accessed June 27, 2024.
39. Hashim M, Pfeiffer BM, Bartsch R, Postma M, Heeg B. Do surrogate endpoints better correlate with overall survival in studies that did not allow for crossover or reported balanced postprogression treatments? An application in advanced non–small cell lung cancer. *Value Health*. 2018;21(1):9–17. doi:10.1016/j.jval.2017.07.011
40. Dreyer NA, Hall M, Christian JB. Modernizing regulatory evidence with trials and real-world studies. *Ther Innov Regul Sci*. 2020;54(5):1112–1115. doi:10.1007/s43441-020-00131-5
41. LoCasale RJ, Pashos CL, Gutierrez B, et al. Bridging the gap between RCTs and RWE through endpoint selection. *Ther Innov Regul Sci*. 2021;55(1):90–96. doi:10.1007/s43441-020-00193-5
42. Liu F, Panagiotakos D. Real-world data: a brief review of the methods, applications, challenges and opportunities. *BMC Med Res Methodol*. 2022;22(1):287. doi:10.1186/s12874-022-01768-6
43. Jaksa A, Wu J, Jónsson P, Eichler HG, Vititoe S, Gatto NM. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. *J Comp Eff Res*. 2021;10(9):711–731. doi:10.2217/cer-2020-0228
44. Zhang J, Symons J, Agapow P, et al. Best practices in the real-world data life cycle. *PLOS Digit Health*. 2022;1(1):e0000003. doi:10.1371/journal.pdig.0000003
45. Liaw ST, Guo JGN, Ansari S, et al. Quality assessment of real-world data repositories across the data life cycle: a literature review. *J Am Med Inform Assoc JAMIA*. 2021;28(7):1591–1599. doi:10.1093/jamia/ocaa340
46. Friends of Cancer Research. Considerations for use of real-world evidence in oncology: a friends of cancer research white paper; 2020. Available from: https://friendsofcancerresearch.org/wp-content/uploads/Use_of_Real-World_Evidence_in_Oncology_0.pdf. Accessed May 3, 2023.
47. Horvat P, Gray CM, Lambova A, et al. Comparing findings from a friends of cancer research exploratory analysis of real-world end points with the cancer analysis system in England. *JCO Clin Cancer Inform*. 2021;5:1155–1168. doi:10.1200/CCI.21.00013
48. Griffith SD, Tucker M, Bowser B, et al. Generating real-world tumor burden endpoints from electronic health record data: comparison of RECIST, radiology-anchored, and clinician-anchored approaches for abstracting real-world progression in non-small cell lung cancer. *Adv Ther*. 2019;36(8):2122–2136. doi:10.1007/s12325-019-00970-1
49. MHRA. MHRA guideline on randomised controlled trials using real-world data to support regulatory decisions. GOV.UK. Available from: https://www.gov.uk/government/publications/mhra-guidance-on-The-use-of-real-world-data-in-clinical-studies-to-support-regulatory-decisions/mhra-guideline-on-randomised-controlled-trials-using-real-world-data-to-support-regulatory-decisions. Accessed May 3, 2022.
50. NICE. Real-world evidence framework feedback. NICE. Available from: https://www.nice.org.uk/about/what-we-do/real-world-evidence-framework-feedback. Accessed May 3, 2022.
51. Health and Social Care, Health and Care Research Wales, NHS Research Scotland, NHS Health Research Authority. Does my project require review by a Research Ethics Committee? Available from: https://www.hra-decisiontools.org.uk/ethics/docs/Algorithm%20-%20Does%20my%20project%20require%20REC%20review%20v2.0%2020200304.pdf. Accessed June 4, 2024.
52. NHS England. National cancer patient experience survey: 2019 survey results. 2019 survey results; 2020. Available from: https://www.ncpes.co.uk/2019-survey-results/. Accessed May 3, 2022.
53. Mohyuddin GR, Koehn K, Abdallah AO, et al. Use of endpoints in multiple myeloma randomized controlled trials over the last 15 years: a systematic review. *Am J Hematol*. 2021;96(6):690–697. doi:10.1002/ajh.26166
54. Hamel JF, Saulnier P, Pe M, et al. A systematic review of the quality of statistical methods employed for analysing quality of life data in cancer randomised controlled trials. *Eur J Cancer Oxf Engl 1990*. 2017;83:166–176. doi:10.1016/j.ejca.2017.06.025
55. Haslam A, Herrera-Perez D, Gill J, Prasad V. Patient experience captured by quality-of-life measurement in oncology clinical trials. *JAMA Network Open*. 2020;3(3):e200363. doi:10.1001/jamanetworkopen.2020.0363
56. Zhu J, Fang F, Sjölander A, Fall K, Adami HO, Valdimarsdóttir U. First-onset mental disorders after cancer diagnosis and cancer-specific mortality: a nationwide cohort study. *Ann Oncol*. 2017;28(8):1964–1969. doi:10.1093/annonc/mdx265
57. Friends of Cancer Research. Establishing a framework to evaluate real-world endpoints; 2018. Available from: https://friendsofcancerresearch.org/wp-content/uploads/RWE_FINAL-7.6.18_1.pdf. Accessed September 21, 2022.
58. Stewart M, Norden AD, Dreyer N, et al. An exploratory analysis of real-world end points for assessing outcomes among immunotherapy-treated patients with advanced non–small-cell lung cancer. *JCO Clin Cancer Inform*. 2019. doi:10.1200/CCI.18.00155
59. Atrash S, Thompson-Leduc P, Tai MH, et al. Treatment patterns and effectiveness of patients with multiple myeloma initiating Daratumumab across different lines of therapy: a real-world chart review study. *BMC Cancer*. 2021;21(1):1207. doi:10.1186/s12885-021-08881-7
60. Akizuki K, Matsuoka H, Toyama T, et al. Real-world data on clinical features, outcomes, and prognostic factors in multiple myeloma from Miyazaki prefecture, Japan. *J Clin Med*. 2020;10(1):105. doi:10.3390/jcm10010105

61. Chari A, Parikh K, Ni Q, Abouzaid S. Treatment patterns and clinical and economic outcomes in patients with newly diagnosed multiple myeloma treated with lenalidomide- and/or bortezomib-containing regimens without stem cell transplant in a real-world setting. *Clin Lymphoma Myeloma Leuk*. 2019;19(10):645–655. doi:10.1016/j.clml.2019.06.007
62. Chen CC, Parikh K, Abouzaid S, et al. Real-world treatment patterns, time to next treatment, and economic outcomes in relapsed or refractory multiple myeloma patients treated with pomalidomide or carfilzomib. *J Manag Care Spec Pharm*. 2017;23(2):236–246. doi:10.18553/jmcp.2017.23.2.236
63. Remes K, Anttila P, Silvennoinen R, et al. Real-world treatment outcomes in multiple myeloma: multicenter registry results from Finland 2009–2013. *PLoS One*. 2018;13(12):e0208507. doi:10.1371/journal.pone.0208507
64. Rice MS, Naeger S, Singh E. Real-world treatment patterns and outcomes among multiple myeloma patients with asthma and COPD in the United States. *Oncol Ther*. 2021;9(1):195–212. doi:10.1007/s40487-021-00146-4
65. Szabo AG, Klausen TW, Levring MB, et al. The real-world outcomes of multiple myeloma patients treated with daratumumab. *PLoS One*. 2021;16(10):e0258487. doi:10.1371/journal.pone.0258487
66. Fiteni F, Westeel V, Bonnetain F. Surrogate endpoints for overall survival in lung cancer trials: a review. *Expert Rev Anticancer Ther*. 2017;17(5):447–454. doi:10.1080/14737140.2017.1316196
67. Avet-Loiseau H, Ludwig H, Landgren O, et al. Minimal Residual Disease Status as a Surrogate Endpoint for Progression-free Survival in Newly Diagnosed Multiple Myeloma Studies: a Meta-analysis. *Clin Lymphoma Myeloma Leuk*. 2020;20(1):e30–e37. doi:10.1016/j.clml.2019.09.622
68. Kothari S, Hillengass J, McCarthy PL, Holstein SA. Determination of minimal residual disease in multiple myeloma: does it matter? *Curr Hematol Malig Rep*. 2019;14(1):39–46. doi:10.1007/s11899-019-0497-7
69. Chen K, Zhao H, Shi Y, et al. Perioperative Dynamic Changes in Circulating Tumor DNA in Patients with Lung Cancer (DYNAMIC). *Clin Cancer Res*. 2019;25(23):7058–7067. doi:10.1158/1078-0432.CCR-19-1213
70. Suzuki K, Nishiwaki K, Gunji T, Katori M, Masuoka H, Yano S. Elevated eosinophil level predicted long time to next treatment in relapsed or refractory myeloma patients treated with lenalidomide. *Cancer Med*. 2020;9(5):1694–1702. doi:10.1002/cam4.2828
71. Provencio M, Serna-Blasco R, Franco F, et al. Analysis of circulating tumour DNA to identify patients with epidermal growth factor receptor–positive non-small cell lung cancer who might benefit from sequential tyrosine kinase inhibitor treatment. *Eur J Cancer*. 2021;149:61–72. doi:10.1016/j.ejca.2021.02.031
72. Charalampous C, Kourelis T. Minimal residual disease assessment in multiple myeloma patients: minimal disease with maximal implications. *Front Oncol*. 2022;11. doi:10.3389/fonc.2021.801851
73. Abbosh C, Birkbak NJ, Swanton C. Early stage NSCLC - challenges to implementing ctDNA-based screening and MRD detection. *Nat Rev Clin Oncol*. 2018;15(9):577–586. doi:10.1038/s41571-018-0058-3
74. Peng Y, Mei W, Ma K, Zeng C. Circulating tumor DNA and Minimal Residual Disease (MRD) in solid tumors: current horizons and future perspectives. *Front Oncol*. 2021;11:763790. doi:10.3389/fonc.2021.763790
75. Kogure Y, Handa H, Ito Y, et al. ctDNA improves prognostic prediction for patients with relapsed/refractory MM receiving ixazomib, lenalidomide, and dexamethasone. *Blood*. 2024;143(23):2401–2413. doi: 10.1182/blood.2023022540
76. Freeman SD, Hills RK, Virgo P, et al. Measurable residual disease at induction redefines partial response in acute myeloid leukemia and stratifies outcomes in patients at standard risk without NPM1 mutations. *J Clin Oncol*. 2018;36(15):1486–1497. doi:10.1200/JCO.2017.76.3425
77. Russell N, Freeman S, Wilhelm-Benartzi C, et al. S135 A single versus fractionated schedule of gemtuzumab ozogamicin in combination with daunorubicin / ARA-C as induction therapy in older patients with aml: results from the UK NCRI AML18 trial. In: *EHA25 Virtual Abstract Book*. HemaSphere. Vol. 4; 2020:19. doi10.1097/HS9.0000000000000404
78. Grimwade D, Jovanovic JV, Hills RK, et al. Prospective minimal residual disease monitoring to predict relapse of acute promyelocytic leukemia and to direct pre-emptive arsenic trioxide therapy. *J Clin Oncol*. 2009;27(22):3650–3658. doi:10.1200/JCO.2008.20.1533
79. Smith G, Apperley J, Milojkovic D, et al. A British Society for Haematology Guideline on the diagnosis and management of chronic myeloid leukaemia. *Br J Haematol*. 2020;191(2):171–193. doi:10.1111/bjh.16971
80. Schievink B, Lambers Heerspink H, Leufkens H, De Zeeuw D, Hoekman J. The use of surrogate endpoints in regulating medicines for cardio-renal disease: opinions of stakeholders. *PLoS One*. 2014;9(9):e108722. doi:10.1371/journal.pone.0108722