Modelling approaches for histology-independent cancer drugs to inform NICE appraisals: a systematic review and decision-framework

Peter Murphy,[1] David Glynn,[2] Sofia Dias,[1] Robert Hodgson,[1] Lindsay Claxton,[1] Lucy Beresford,[1] Katy Cooper,[3] Paul Tappenden,[3] Kate Ennis,[3] Alessandro Grosso,[2] Kath Wright,[1] Anna Cantrell,[3] Matt Stevenson[3] and Stephen Palmer[2]*

[1]Centre for Reviews and Dissemination, University of York, York, UK
[2]Centre for Health Economics, University of York, York, UK
[3]School of Health and Related Research (ScHARR) Technology Assessment Group, University of Sheffield, Sheffield, UK

*Corresponding author stephen.palmer@york.ac.uk

Declared competing interests of authors: none.

Published December 2021
DOI: 10.3310/hta25760

Scientific summary

Modelling approaches for histology-independent cancer drugs
Health Technology Assessment 2021; Vol. 25: No. 76
DOI: 10.3310/hta25760

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

In 2017, the US Food and Drug Administration (FDA) granted approval to pembrolizumab (Keytruda, Merck Group, Darmstadt, Germany) for the treatment of solid tumours with the microsatellite instability high (MSI-H) or the deficient mismatch repair (dMMR) biomarker. This was the first time that a cancer treatment was approved based on a common biomarker rather than the location in the body at which the tumour originated. It represented an important paradigm shift, which means that oncological diseases can now be classified by either tumour biomarker status or tumour histogenesis. The first histology-independent marketing authorisation was granted by the European Medicines Agency (EMA) in 2019.

A histology-independent marketing authorisation will include a large number of individual tumour sites. It is unlikely to be feasible or desirable for the National Institute for Health and Care Excellence (NICE) to conduct separate appraisals for each individual tumour site to inform whether or not approval of these products represents an efficient use of NHS resources. However, the scope of histology-independent indications and the nature of the evidence base will pose important challenges to the appropriate quantification of their value and the effective mitigation of any additional risks. NICE needs to consider how to develop a process that will allow a single, biomarker-driven appraisal for these drugs.

This research aims to inform future NICE policy on how to appraise cancer drugs with histology-independent indications.

Objectives

We sought to explore the implications of histology-independent products for the NICE technology appraisal (TA) process. The specific objectives were to:

1. identify the nature of the evidence likely to be available at initial marketing authorisation
2. determine the types of evidence and analyses required to support NICE appraisal
3. develop a case study to highlight methods and evidence challenges, and to explore alternative ways of addressing these challenges
4. develop a conceptual framework to establish the evidence and analyses required to guide NICE decision-making and potential Cancer Drugs Fund (CDF) data collection requirements
5. suggest any changes to the current NICE methods guide or additional requirements relating to histology-independent drugs
6. make recommendations for further research.

Methods

We undertook a series of targeted reviews to determine the type of evidence that is likely to be available at initial marketing authorisation and to consider the analyses required to support a NICE appraisal. These reviews included:

1. a review of FDA and EMA websites to identify relevant documents relating to regulatory issues and benefit–risk approaches for histology-independent indications
2. an overview of key statistical literature addressing the design and analysis of histology-independent trials
3. a systematic review to identify published meta-analyses evaluating the use of overall response rate (ORR) and duration of response (DoR) as surrogate end points for progression-free survival (PFS) and overall survival (OS).

These reviews were used to identify specific challenges for histology-independent appraisals and to identify approaches that might be used to investigate and account for different sources of uncertainty and heterogeneity.

We developed an exemplar model to illustrate the nature of the assessments that could be used to assess the cost-effectiveness of a new histology-independent treatment and to inform NICE decision-making. A framework to inform approval and research policies was also proposed to help determine the appropriateness of different policy recommendations and to identify key uncertainties that might inform and prioritise the value of further data collection.

Based on these findings, a series of recommendations were made concerning potential changes to the current NICE methods guide and priorities for further methodological research.

Results

Review of regulatory issues and benefit–risk approaches relevant for histology-independent indications

Our review found that histology-independent products are likely to be evaluated using more complex study designs that are intended to increase the efficiency of the drug development process, specifically basket trials with master protocols. Master protocols are used to evaluate multiple drugs and/or multiple cancer subpopulations in parallel, using a single protocol. Basket trials are used to evaluate a single investigational drug or drug combination in different populations (defined by disease stage, histology or treatment history), and are usually designed as single-arm activity-estimating trials with ORR as the primary end point.

Our review highlighted the importance placed by the regulators in the underlying biological rationale and strength of existing clinical evidence to support the assumption that a biomarker-defined population is sufficient to establish clinically relevant activity, independent of tumour histology. Importantly, neither the FDA nor the EMA concluded that the evidence for the existing histology-independent products was sufficient to support a routine approval decision. Although the treatment effect observed in the overall population was considered to be clinically important, the initial approvals were conditional on additional requirements for further evidence generation to increase the precision of the effect estimates and extend the length of follow-up. Hence, important new evidence will emerge over time.

Overview of key statistical literature addressing the design and analysis of trials

A critical consideration in the design of histology-independent trials is the potential for heterogeneity in prognosis across the different histologies; therefore, standardised response rates, reflecting tumour shrinkage, are typically used instead of survival outcomes. In addition, randomisation to a control arm is rare in basket trials owing to the differences in standard of care across the different tumour types. The reliance on surrogate outcomes and the lack of a concurrent, randomised, control arm remains a key limitation of these trial designs and for the interpretation of such trials for NICE appraisal.

Heterogeneity of effect across different baskets is a key consideration in the analysis of histology-independent trials. Once a decision has been made on whether or not heterogeneity is present, the analysis typically proceeds either as separate, independent studies for each basket or as a single aggregate study combining all of the baskets. Thus, either complete homogeneity or completely unrelated effects are assumed. A less restrictive assumption is that efficacy is similar (rather than equal or completely different) across baskets. Bayesian hierarchical models (BHMs) are particularly suited for this situation because they estimate the heterogeneity and allow borrowing of information on the

effects of the treatment across baskets. However, the BHM is advantageous only if it is considered reasonable to allow such borrowing. Alternatives to complete pooling or borrowing have been proposed, which extend the BHM to allow borrowing of information across similar baskets while avoiding overly optimistic borrowing for extreme baskets.

Although it is challenging to determine the correct level of borrowing of information, BHM approaches provide an explicit basis to allow the treatment effect in any basket to be informed by the effects in all other baskets, therefore maximising the information available.

A systematic review to identify published meta-analyses evaluating the use of response rates and duration of response as surrogate end points for progression-free and overall survival

In the context of histology-independent treatments, data on OS and potentially other time-to-event outcomes, such as PFS, are likely to be immature. Consequently, there may be a need to rely on surrogate outcomes, such as response rate, using data from external sources to estimate other more clinically meaningful final outcomes for NICE appraisal. We undertook a systematic review to assess the strength of the association between response outcomes and PFS, time to progression (TTP) or OS across different types of cancer (primarily advanced or metastatic), based on meta-analyses or meta-regression studies assessing the statistical relationship between these outcomes. Alternative sets of criteria were used to assess the strength of association between surrogate and final end points.

A total of 63 studies were included in the review, across 20 different cancer types. The most commonly analysed relationships were between ORR and either PFS or OS. The association between response outcomes and PFS/TTP/OS was found to vary widely between studies, and generally scored low to medium when assessed using existing criteria. No clear pattern for strength of association was identified by cancer type.

Our findings indicate that response end points may not be reliable surrogates for PFS or OS. However, despite the weak validity of response as a surrogate for PFS and OS, we concluded that it might still be preferable for NICE appraisals to adopt a surrogate-based modelling approach informed by predictions from meta-analyses that capture all relevant uncertainty, rather than ignoring surrogate relationships and extrapolating heavily censored PFS and OS data.

Exemplar case study

We also identified several additional potential challenges for NICE appraisals, including the need to account for heterogeneity in a number of areas, including the cost of testing, baseline risk, quality of life and routine management costs. A range of alternative analytical approaches are likely to be required to address these different areas.

The use of a single assessment of the incremental cost-effectiveness ratio (ICER) across multiple tumour sites with potentially different treatment effectiveness, comparators, costs and quality of life may be challenging for NICE to interpret. A single ICER may conceal significant variation in the tumour-specific ICERs, driven by a combination of factors, including observable variability in relative effectiveness between tumour types. Ignoring these differences could mean that a treatment that is not cost-effective for the total population may be cost-effective in specific subgroups. Conversely, a treatment that appears cost-effective for the total population may not be cost-effective in particular subgroups. Given the amount of heterogeneity associated with a histology-independent appraisal, estimating the average cost-effectiveness for the full patient population covered by the product's licence may not provide enough information to decision-makers about whether or not the drug is potentially cost-effective across all subgroups.

Given the importance of exploring the impact of heterogeneity on decision-making, explicit and transparent approaches are required that can accommodate different sources of heterogeneity within the overall population. These assessments should allow consideration of the average cost-effectiveness for the full patient population covered in the marketing authorisation, as well as facilitating an assessment of whether or not the drug is potentially cost-effective across relevant subgroups. The BHM framework provides an important approach that can more fully explore the potential heterogeneity in effects across tumours. The BHM approach allows assessments to be made for each tumour type, as well as a pooled assessment across all tumour types, accounting for the potential lack of uniformity of effect across tumours. An additional advantage of this type of model is the ability to predict the response probability that would be expected in a 'new' tumour type (i.e. a tumour that is not represented in the trial data), which will give a measure of the uncertainty in the response rates in tumour types in the target population but for which no data are available.

An exemplar case study was developed to illustrate the nature of the assessments that could be used to evaluate the cost-effectiveness of a new histology-independent treatment. The case study considered a hypothetical tyrosine kinase (TRK) inhibitor ('Drug X') for the treatment of solid tumours that harbour a neurotrophic tyrosine receptor kinase (*NTRK*) gene fusion compared with the current standard of care. The economic model used a landmark response-based structure incorporating separate PFS and OS distributions, conditioned on response status in the overall study population. Heterogeneity in response rates across individual tumour sites was reflected using a BHM approach. By linking the BHM estimates for response rates to conditional OS and PFS estimates, the case study model explores the implications for cost-effectiveness of heterogeneity in the overall population by considering individual histology-specific estimates of cost-effectiveness alongside estimates for the overall population.

In line with the NICE reference case, the model was based on a NHS and Personal Social Services perspective using a 3.5% discount rate. Results are presented over a lifetime (i.e. 30-year) time horizon.

The case study demonstrated the importance of understanding the frequency of histologies expected in the target population and the necessity of modelling histology-specific costs and health consequences. When the expected distribution of histologies is expected to differ between the trial and the target population, failure to account for this could result in a biased estimate of the pooled ICER. The magnitude of any bias will depend on the extent of heterogeneity in the relevant model inputs between tumour sites. Consideration will also be needed as to the potential effect in tumour histologies that are not represented in the trial data.

The case study also highlighted that even if homogeneity in all other model inputs is assumed between individual histologies (or other subgroups), the cost-effectiveness estimates will inevitably vary based on differences in the costs of identifying patients with the specific biomarker. The results demonstrate that even a low per-patient testing cost can result in significant variation in the ICER estimates driven by different biomarker prevalence rates across individual histologies.

The case study was used to illustrate how heterogeneity could be explored using pooled ICERs and individual histology ICERs to inform decision-making. However, ICERs have an important limitation: they do not give an indication of the scale of consequences for population health. Understanding the benefits and costs of treatment at a population level will help to interpret the consequences of decision-making in the presence of heterogeneity and uncertainty.

We developed a decision framework that could be used to inform approval and research policies for histology-independent products. The framework explored the uncertainties and risks associated with different approval policies. Alternative approaches to managing risk were identified, including the role of further data collection, the use of pricing schemes and stratified decision-making.

Conclusions

Our research found that the potential for heterogeneity in a range of model inputs, either across tumour histologies or other characteristics, is likely to be an important issue for NICE appraisals of histology-independent technologies. Consideration should be given to the appropriateness of the assumptions of homogeneity of treatment effects and NICE committees should expect to see an exploration of this assumption in company submissions. Where there is evidence of heterogeneity in treatment effects and estimates of cost-effectiveness, consideration should be given to optimised or 'stratified' recommendations. Routine presentation of the scale of the consequences of heterogeneity and decision uncertainty may provide an important additional approach to the assessments specified in the current NICE methods guide.

We identified several areas requiring further research. First, further exploration of BHM could help to determine whether or not there is sufficient evidence of homogeneity to support pooled analyses. Second, further research is required to determine the appropriate basis for apportioning genomic testing costs where there are multiple targets. Finally, further research is required concerning the challenges of uncontrolled Phase II studies and specifically the role and use of surrogate end points.

Funding

This project was funded by the National Institute for Health Research (NIHR) Evidence Synthesis programme and will be published in full in *Health Technology Assessment*; Vol. 25, No. 76. See the NIHR Journals Library website for further project information.

Health Technology Assessment

HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.014

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, the Cochrane Library and Clarivate Analytics Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

This report

The research reported in this issue of the journal was commissioned and funded by the Evidence Synthesis Programme on behalf of NICE as project number NIHR127852. The contractual start date was in November 2018. The draft report began editorial review in March 2020 and was accepted for publication in May 2021. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health and Social Care.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).