PharmacoEconomics (2017) 35:1287–1296
DOI 10.1007/s40273-017-0562-9

ORIGINAL RESEARCH ARTICLE

The HTA Risk Analysis Chart: Visualising the Need for and Potential Value of Managed Entry Agreements in Health Technology Assessment

Sabine Elisabeth Grimm[1] · Mark Strong[2] · Alan Brennan[2] · Allan J. Wailoo[2]

Published online: 28 August 2017
© The Author(s) 2017. This article is an open access publication

Abstract

Background Recent changes to the regulatory landscape of pharmaceuticals may sometimes require reimbursement authorities to issue guidance on technologies that have a less mature evidence base. Decision makers need to be aware of risks associated with such health technology assessment (HTA) decisions and the potential to manage this risk through managed entry agreements (MEAs).

Objective This work develops methods for quantifying risk associated with specific MEAs and for clearly communicating this to decision makers.

Methods We develop the 'HTA risk analysis chart', in which we present the payer strategy and uncertainty burden (P-SUB) as a measure of overall risk. The P-SUB consists of the payer uncertainty burden (PUB), the risk stemming from decision uncertainty as to which is the truly optimal technology from the relevant set of technologies, and the payer strategy burden (PSB), the additional risk of approving a technology that is not expected to be optimal. We demonstrate the approach using three recent technology appraisals from the UK National Institute for Health and Clinical Excellence (NICE), each of which considered a price-based MEA.

Results The HTA risk analysis chart was calculated using results from standard probabilistic sensitivity analyses. In all three HTAs, the new interventions were associated with substantial risk as measured by the P-SUB. For one of these technologies, the P-SUB was reduced to zero with the proposed price reduction, making this intervention cost effective with near complete certainty. For the other two, the risk reduced substantially with a much reduced PSB and a slightly increased PUB.

Conclusions The HTA risk analysis chart shows the risk that the healthcare payer incurs under unresolved decision uncertainty and when considering recommending a technology that is not expected to be optimal given current evidence. This allows the simultaneous consideration of financial and data-collection MEA schemes in an easily understood format. The use of HTA risk analysis charts will help to ensure that MEAs are considered within a standard utility-maximising health economic decision-making framework.

Key Points for the Decision Maker

The health technology assessment (HTA) risk analysis chart presents a standardised visualisation to show the need for and potential value of different classes of managed entry agreement (MEA) schemes.

Its use in HTA could ensure that MEAs are considered routinely, consistently and transparently.

The HTA risk analysis chart allows for simultaneous consideration of financial and data-collection MEA schemes.

1 Introduction

Recent changes to the regulatory landscape of pharmaceuticals, such as adaptive pathways or conditional licensing schemes [1, 2] issued by the European Medicines

✉ Sabine Elisabeth Grimm
sabine.grimm@mumc.nl

[1] Maastricht University Medical Center, Maastricht, The Netherlands

[2] School of Health And Related Research, University of Sheffield, Sheffield, UK

Agency, allow licensing of new health technologies more quickly, or with a restriction to a certain population. These changes result in decision-making bodies being required to issue guidance on technologies that have a smaller evidence base than previously, causing greater uncertainty regarding the clinical and cost effectiveness of new technologies at the point of decision making. With this comes an increased risk of making the 'wrong' decision that affects the healthcare system and ultimately health in the population served.

Schemes that allow the development of further evidence or that entail a risk-sharing component can be employed to mitigate this risk. These schemes, called managed entry agreements (MEAs) [3], are agreements between manufacturers and decision-making bodies designed to reduce the risk incurred by health services. The two broad conditions set out in MEAs are (1) that the price of the technology be reduced through a range of different financial schemes and/or (2) that further data will be collected [3, 4]. Both conditions aim to reduce the risk associated with a recommendation and include a range of specific types of schemes that are described in detail in the literature [3–7]. A recent review of MEAs, or performance-based risk-sharing schemes, showed that companies are often willing to share risk in schemes that entail coverage with evidence development, price reductions or performance-linked reimbursement, conditional treatment continuation, and financial or utilisation caps as well as in multi-component schemes [8]. The review furthermore found that the use of such schemes is increasing globally [8]. However, the exact conditions can vary widely in terms of their duration, populations in which they are employed, stopping rules and other features [8].

Health technology assessment (HTA) decision making follows the principle that when a new technology is estimated to be cost ineffective, its recommendation is not warranted at the given price. However, in practice, decision makers are faced with decisions under uncertainty and a range of plausible incremental cost-effectiveness ratios (ICERs) that often span cost-effectiveness thresholds. Whilst evidence collection-based MEAs will reduce uncertainty and produce revised estimates of costs and benefits after data are collected, this does not alter the calculation of expected net benefit of the appraised technologies at present, that is before the evidence has been collected. Similarly, a financial MEA scheme will not directly address the scale of uncertainty in costs or benefits. However, price reductions can and do affect decision uncertainty by making a technology more (or less) likely to be cost effective given current evidence. There is therefore value in an analysis that considers either a price reduction or a research-based MEA scheme, and indeed one that can consider schemes that incorporate both reducing uncertainty and price simultaneously.

The existing health economic literature provides methods for assessing research schemes [4, 9–13]. Furthermore, the requirements that have to be in place for 'only in research' or 'recommendation with research' (RwR) schemes, and the principles and assessment methods for such schemes, have been discussed in detail by Claxton et al. [12], McKenna et al. [9] and Rothery et al. [5]. These papers are particularly useful in deciding whether more research could and should be conducted, and in which sequence different assessments should be performed. However, these papers do not provide a method for quantifying and presenting risks that are typically encountered at the time of decision making. Since the above-mentioned trend of an increasing number of company submissions with an immature evidence base may result in decision makers demanding more research in many cases, we consider it important to have a unified approach for simultaneously considering the need for price and research MEA schemes in the HTA process and to be able to present the results of such risk analyses to analysts and decision makers in an accessible and comprehensible manner.

This paper describes and presents a visual aid for decision making that enables the simultaneous consideration of financial and research schemes—the HTA risk analysis chart. This builds upon existing methods referenced above and enables consistent and transparent consideration of the potential need for MEA schemes in the HTA process. The paper is structured as follows: in the Methods section, we introduce the key concepts of the 'payer strategy and uncertainty burden' (P-SUB) to assess the need for an MEA, we develop the HTA risk analysis chart and describe three past UK National Institute for Health and Care Excellence (NICE) technology appraisals in which we applied it; in the Results section, we present the HTA risk analysis chart applied to the three case studies, discuss the potential need for an MEA in each of these appraisals and present the change in the HTA risk analysis charts after an MEA scheme is put in place; we end with a discussion and conclusion.

2 Methods

2.1 The Payer Strategy and Uncertainty Burden as a measure of risk in Health Technology Assessments

We consider two types of risk burden linked with a decision on whether to recommend the introduction of a new health technology in a health system. The first is the risk burden that arises due to decision uncertainty and is a characteristic that applies to all the technologies or strategies that are compared in the overall decision problem. The

second is the strategy-specific risk burden associated with choosing a specific non-optimal strategy from the set of technologies available.

Our context is a decision problem with at least two technologies compared, where the decision analyst calculates the expected costs and benefits of each technology given current evidence and proposed prices, accounting for uncertainty. The decision maker then determines which of the technologies should be considered 'optimal'. The most optimal strategy is that with the greatest expected 'payoff' as measured by expected net benefit. The first risk that we consider is based on decision uncertainty and can be more intuitively described as the risk that the health technology that appears optimal based on current evidence and prices might, in fact, not be the truly optimal strategy. This may be caused by uncertainty in cost-effectiveness model parameters due to imperfect current evidence.

The second risk we consider is that associated with adopting a 'non-optimal' technology, i.e. one that appears, given current evidence and current proposed prices, to be less cost effective than the optimal technology. As this 'optimality' is measured by expected net benefit, this concept simply refers to the risk of adopting a technology that does not have the highest expected net benefit. Generally, a decision maker would recommend the most cost-effective strategy (based on decision theory) unless risk neutrality is violated. This second risk therefore answers the what-if question: "What if a 'non-optimal' technology were recommended—how much loss or burden would that incur (again using the expected net benefit measure)?". This may be important when the decision maker does not recommend the optimal technology, because of a deviation from risk neutrality. A rational decision maker, if risk neutral, and under the assumption that any decision is reversible at no cost, should always select the decision option with the highest expected net benefit. They should never select an option with a positive payer strategy burden (PSB). However, when faced with highly uncertain decisions, decision makers (who are not necessarily risk neutral) may be inclined to consider recommending technologies expected to be cost ineffective, under the condition that further research is undertaken. This second risk measure is therefore important, and it is strategy specific—for the strategy that is expected to be optimal, the value of this risk burden is zero, but for each 'non-optimal' strategy, the value of this risk burden will be positive.

We propose quantifying and visualising the first risk, i.e. the measure of the overall risk burden associated with all possible decisions, by using the payer uncertainty burden (PUB). The PUB describes the expected cost of decision uncertainty, which is related to the probability of making the 'wrong' decision based on current evidence, and the cost associated with that 'wrong' decision. The PUB is equivalent to the expected value of perfect information (EVPI), or the expected opportunity loss [14], and can be interpreted as the value of eliminating all uncertainty and hence the possibility of making the wrong decision [15]. The calculation of the PUB, or the EVPI, is based upon results from the probabilistic sensitivity analysis (PSA), and was intuitively described by Wilson [16]. Our rationale for introducing the new term 'payer uncertainty burden' for EVPI was to make it absolutely clear that it is not only the uncertainty itself that is quantified but also the consequences of this uncertainty, which pose a burden to the payer, compared with a scenario in which there is no uncertainty. Whilst we do not wish to replace the term EVPI, we do wish to spark a change in thinking to reflect that the PUB is not only the value of research but a risk burden to the decision maker and payer (originally referred to as the expected opportunity loss by Raiffa in 1968 [14]), which can then be addressed through the use of MEAs. The PUB therefore describes the risk to the payer (on behalf of society), and, in a health system with a budget constraint, ultimately the expected risk to patients in terms of societal health foregone. Mathematically, this can be written as

$$\begin{aligned}\text{PUB} &= \text{EVPI} \\ &= \left[\mathbb{E}_\theta\left\{\max_d \text{NB}(d,\theta)\right\} - \max_d \mathbb{E}_\theta\{\text{NB}(d,\theta)\}\right] \geq 0,\end{aligned} \tag{1}$$

where $\text{NB}(d,\theta)$ is the net benefit function, d indexes strategies in some set D, and θ is a vector of uncertain model parameters.

We propose quantifying and visualising the second risk, i.e. the measure of the strategy-specific risk burden, given current evidence and price, via a second quantity, the payer strategy burden (PSB). The PSB for each specific strategy is the difference between the expected net benefit of the optimal strategy and the expected net benefit of the chosen strategy. Again, mathematically,

$$\text{PSB}(d') = \left[\max_d \mathbb{E}_\theta\{\text{NB}(d,\theta)\} - \mathbb{E}_\theta\{\text{NB}(d',\theta)\}\right] > 0, \tag{2}$$

where d' is a strategy that is expected to be cost ineffective based on current evidence. For the cost-effective strategy d^*, the $\text{PSB}(d^*)$ equals zero.

If in a hypothetical technology assessment, a technology expected to be cost ineffective was recommended under decision uncertainty, the payer would face the combined risk of the PSB and the PUB. We denote the sum of these as the P-SUB. Each of these quantities (PUB, PSB, P-SUB) can be expressed in either monetary units or health output units (for example, life-years or quality-adjusted life-years [QALYs]). The $\text{P-SUB}(d')$ is given by,

$$\begin{aligned}\mathrm{PSUB}(d') &= \mathrm{PUB} + \mathrm{PSB}(d') \\ &= \left[\mathbb{E}_\theta\left\{\max_d\{\mathrm{NB}(d,\theta)\}\right\} - \mathbb{E}_\theta\{\mathrm{NB}(d',\theta)\}\right].\end{aligned} \tag{3}$$

Given the size of the population for whom the decision problem is relevant in a year, the annual population PUB, PSB or P-SUB can be calculated. This allows the proposed risk measures to be compared in population-level absolute terms across decision problems.

2.2 The HTA Risk Analysis Chart

We propose the HTA risk analysis chart as a method for immediately conveying the P-SUB associated with each strategy in a decision problem in a single, simple plot. An example for an illustrative model is shown in Fig. 1. The blue bars represent the PUB and are the same height for each intervention because the PUB is the risk relating to uncertainty associated with the whole decision problem rather than any specific decision strategy. The overall EVPI, i.e. the PUB, is £700 per person affected by the decision, which at a maximum acceptable ICER of £20,000 is equivalent to 0.035 QALYs worth of decision uncertainty per person. The PSB is represented by the red bars stacked on top of the PUB. These are different heights because they relate to the strategy-specific risk. The cost-effective intervention (in this figure, intervention 3) has a PSB of zero. Given current prices and evidence, both intervention 2 and intervention 3 are less cost effective than intervention 1, which is indicated by their respective PSBs of £1000 and £2000 per person. The P-SUBs are shown on the cost scale on the *y* axis (£1700 and £2700, respectively) and on the QALY scale above each bar (0.085 and 0.135 QALYs, respectively). We also present in a text box the PUB (£7 million) and the largest PSB (£20 million) of the decision problem accrued over the affected patient population per annum. This enables cross comparison between decision problems in terms of the national scale of risk involved.

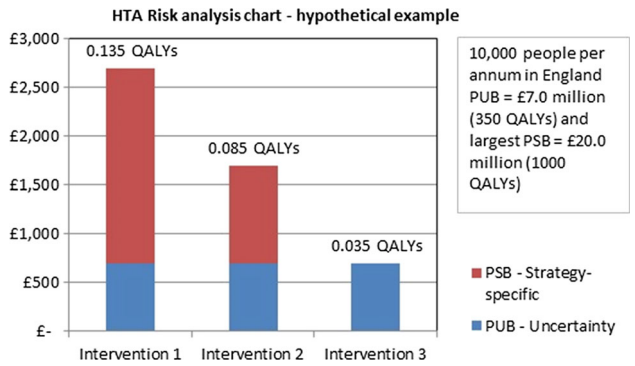

Fig. 1 Health technology assessment risk analysis chart illustrated in a hypothetical example. *PSB* payer strategy burden, *PUB* payer uncertainty burden, *QALY* quality-adjusted life-year

2.3 Using the HTA Risk Analysis Chart to Consider Managed Entry Agreements

The quantified P-SUB provides information that can help to inform the need for and potential value of an MEA, as well as the form of MEA (whether price based and/or evidence based).

The first step is to examine the risk given current evidence and proposed prices, as for example in our hypothetical example in Fig. 1. If there is a substantial PSB (a large red component of the bar) for the intervention of interest, this suggests that a price-based MEA could be useful. This is because a price reduction in the technology of interest would decrease the PSB, i.e. decrease the expected incremental net benefit between the technology of interest and the technology we currently expect to be most cost effective. If there is a large PUB (a large blue component to each bar) due to uncertainty in model parameters based on current evidence, this suggests a further evidence collection-based MEA scheme could warrant further investigation [4]. The PUB is equivalent to the EVPI, and an additional step that can be useful to help design proposed evidence collection-based MEAs is to establish which of the uncertain parameters is driving decision uncertainty by performing an expected value of perfect parameter information (EVPPI) analysis as described elsewhere [17, 18]. Whilst these rules of thumb are broadly true, it is important to realise that either type of MEA scheme (price based or evidence based) has an effect on both the PSB and the PUB, i.e. any MEA scheme will change the size of both the red and the blue bars in the HTA risk analysis chart.

The second step is to estimate revised risk analysis charts by simulating proposed MEA schemes. The analysis required depends upon the type of MEA scheme proposed. The process for calculating the revised P-SUB is simpler for price reductions than for research schemes, although both follow the same principles. The easiest type of scheme to analyse is a simple proposed price discount. For this, the cost-effectiveness model can simply be re-run with the new price in place and the resulting PSA can be used to produce the revised HTA risk analysis chart. More complicated price-based MEA schemes exist where price is contingent upon health outcomes, for example a money-back guarantee scheme [3], where the payer only pays for the treatment for those patients who experience a response above some pre-set threshold. For such price-reduction schemes, it is necessary to perform another PSA with the MEA price rule in place.

In contrast to price-reduction schemes, additional steps are required for assessing evidence-based schemes: (1) the

calculation of EVPPI to obtain an idea of where future research efforts should be directed; (2) the design of potential research schemes, for example with different sample sizes and follow-up; (3) the calculation of the expected value of sample information (EVSI) of these schemes; (4) the comparison of the different EVSIs to costs by subtracting the costs of each scheme from its EVSI to yield the expected net gain of this research [11]; and (5) the time that elapses between the decision and the time at which research becomes available, takes an effect and translates into a gain in population health. The EVSI quantifies the expected reduction in uncertainty if we are to undertake a specific data-collection exercise as part of an MEA scheme, for example a new randomised controlled trial (RCT) or a post-marketing observational study [10, 13, 19]. It turns out, mathematically, that the reduction in the PUB per person due to a proposed evidence-based scheme is exactly the EVSI, so it is simple to calculate the revised risk analysis chart once the EVSI calculation has been undertaken. The mathematics of this [6] and the methods to calculate EVSI, including fast estimation of EVSI using the SAVI online tool, are described in further detail elsewhere [6, 10, 13]. It may be of interest to note that calculation of EVPPI or EVSI can only give an indication of the expected with-research P-SUB, but once the data have been collected, the PUB can be recalculated based on the new evidence. The new PUB will have reduced, assuming the study was designed to address the present uncertainties and was well designed and executed. The PSB may also have changed, for instance if the research finds that the new technology is more or less effective than previously thought.

We define the revised P-SUB as that which remains when an MEA is adopted. The overall value of any proposed MEA design can be assessed in terms of its reduction in the P-SUB, i.e. the original P-SUB given current evidence and proposed prices minus the revised P-SUB with the MEA scheme in place. When assessing MEA schemes that take effect only at future time points (most research schemes), it is also desirable to present the P-SUB over the expected technology relevance horizon. The P-SUB should then be calculated for each time period, with and without the use of an MEA scheme, to compare the lifetime value of that MEA scheme with the counterfactual. Further detail on this can be found in Grimm et al. [6].

2.4 Application of the HTA Risk Analysis Chart in Three Technology Appraisals from the UK National Institute for Health and Care Excellence

We illustrate the HTA risk analysis chart using three existing technology appraisals conducted by NICE. Each case involved decision uncertainty and a financial MEA proposal. We had access to the cost-effectiveness models for each of the three appraisals and were therefore able to calculate the P-SUB with and without the financial MEA in place.

2.4.1 Study 1

In 2010, NICE appraised trabectedin versus best supportive care for the treatment of advanced soft tissue sarcoma [20]. Trabectedin was recommended in patients for whom treatment with anthracyclines and ifosfamide had failed or in patients who were intolerant of or had contraindications for treatment with anthracyclines and ifosfamide under the conditions of a financial MEA. This MEA entailed that the acquisition cost of trabectedin for treatment needed after the fifth cycle be met by the manufacturer. Trabectedin was considered to fulfil end-of-life criteria, i.e. to extend life by more than 3 months for patients with a life expectancy shorter than 24 months [21]. To reflect the additional value placed on such end-of-life technologies in the UK, the maximum acceptable ICER used in the present analysis was £50,000 per QALY gained. To reflect this proposed MEA scheme, the model was adjusted to incorporate the 'no reimbursement for more than five treatment cycles' rule.

2.4.2 Study 2

In 2012, NICE appraised dasatinib and nilotinib versus interferon-α for the treatment of imatinib-resistant chronic myeloid leukaemia (CML) [22], amongst other treatments and indications that, for the sake of simplicity, are not presented here. The manufacturers of nilotinib and dasatinib had agreed to patient access schemes that entailed straight discounts on the list price. NICE recommended nilotinib but did not recommend dasatinib. The maximum acceptable ICER used in the present analysis was £20,000 and £30,000 per QALY gained. To adjust the PSA to reflect the proposed MEA schemes, we used the discounted price instead of the list price.

2.4.3 Study 3

In 2014, NICE appraised lenalidomide versus best supportive care for the treatment of myelodysplastic syndromes [23]. It was recommended as an option for people with transfusion-dependent anaemia caused by low- or intermediate-risk myelodysplastic syndromes associated with an isolated deletion 5q cytogenetic abnormality when other treatments have failed to achieve a full response. Approval was conditional on a financial MEA scheme. The MEA scheme was a utilisation cap scheme by which the

manufacturer would not be reimbursed for more than 26 treatment cycles per patient. The maximum acceptable ICER used in the present analysis was £20,000 per QALY gained. The model was adjusted to reflect this proposed MEA scheme by capping the cost at a maximum of 26 treatment cycles.

3 Results

In study 1, trabectedin, when offered at list price, was associated with an ICER of £52,000 per QALY gained, a PUB of £1457 and a PSB of £1000 per person (assuming a £50,000 per QALY threshold). The P-SUB was therefore £2457 per person, or £2.4 million for the 1000 annually affected patient population (Fig. 2a). The proposed financial MEA scheme made trabectedin cost effective (ICER of £35,000 per QALY gained), i.e. reduced the PSB to zero, and also eliminated any decision uncertainty, i.e. showing a PUB of zero and therefore a P-SUB of zero (Fig. 2b). The elimination of the PUB was not caused by a reduction in uncertainty in the cost or benefit parameters. Instead, the price scheme eliminated decision uncertainty by moving the ICER away from the threshold. The risk reduction achieved due to the MEA as measured by the change in P-SUB at a population level was therefore around £2.4 million. With a P-SUB of zero, no further MEA schemes were required.

Without the MEA, dasatinib and nilotinib in study 2 had ICERs of approximately £25,000 per QALY gained compared with interferon-α, and both were associated with a large PSB (assuming a threshold of £20,000 per QALY gained) of £24,300 and £17,200 per person, respectively (Fig. 2c). With 200 people affected annually in the population, these are equivalent to £4.86 million and £3.44 million. The PUB was only £8 per person, reflecting small uncertainty about these interventions being cost ineffective when compared with interferon-α. The proposed MEAs for nilotinib and dasatinib resulted in only a small change in the ICERs and only marginally changed the HTA risk analysis chart. However, when a threshold of £30,000 per QALY gained was considered, nilotinib had a PSB of zero, but this also resulted in an increase in decision uncertainty, reflected in a post-MEA PUB of £1500 per person (Fig. 2d). The MEA scheme for dasatinib reduced its PSB to £5500 per person. Further price reductions through financial schemes would be needed to eliminate the PSB for dasatinib. The PUB accrued over the population was not large, at £0.3 million, making it unlikely though not impossible that further research could have a positive expected net benefit. For nilotinib, the risk reduction achieved due to the MEA as measured by the change in P-SUB at a population level was around £3.1 million, whereas that for dasatinib was around £3.46 million.

In study 3 and without the MEA, lenalidomide was cost ineffective (ICER of £70,000 per QALY gained, and a threshold of £20,000 per QALY gained was assumed) compared with best supportive care. The PSB was large at just under £35,000 per person and the PUB was zero, reflecting that there was no uncertainty about lenalidomide being cost ineffective (Fig. 2e). The large per person PSB also translated into a large impact accrued over the affected patient population, as the payer would incur a PSB of almost £250 million, which would translate into approximately 12,500 QALYs foregone. With the MEA, lenalidomide remained cost ineffective (ICER of £25,000 per QALY gained), but the PSB was substantially reduced to £4700 per person (Fig. 2f), or £26 million when accrued over the affected population. The MEA did also introduce some decision uncertainty, reflected in a PUB of £1100 per person, or £8.5 million accrued over the affected population. Therefore, for lenalidomide, the risk reduction achieved due to the MEA as measured by the change in P-SUB at a population level was around £215 million. The magnitude of the residual P-SUB (£34.5 million) indicates that there would still be potential value in further research or financial MEAs.

Comparison across the three technology appraisals is facilitated by the HTA risk analysis chart. It is easy to see, visually and with the summary measures, that the risk prior to the proposed MEA is substantially larger in the lenalidomide (Fig. 2e) than in the trabectedin appraisal (Fig. 2a), with the risk for dasatinib and nilotinib (Fig. 2c) in between those two. When multiplying up by the much larger population affected for lenalidomide, the pre-MEA population-level risk burden is around 50 times greater for this technology than for any others examined in our studies (£249 million population P-SUB vs. £4.9 million for the next largest).

4 Discussion

We have presented the HTA risk analysis chart as an approach to visualising the need for and potential value of MEA schemes in a consistent and transparent manner, building on standard methods already used in HTA. We developed the concept of P-SUB to assess the risk associated with HTA decisions. This measure enables the simultaneous consideration of decision uncertainty and the extent to which a technology is expected to be cost ineffective. MEA schemes can then be assessed using the reduction in the P-SUB they can achieve. Three examples based on past NICE technologies were presented.

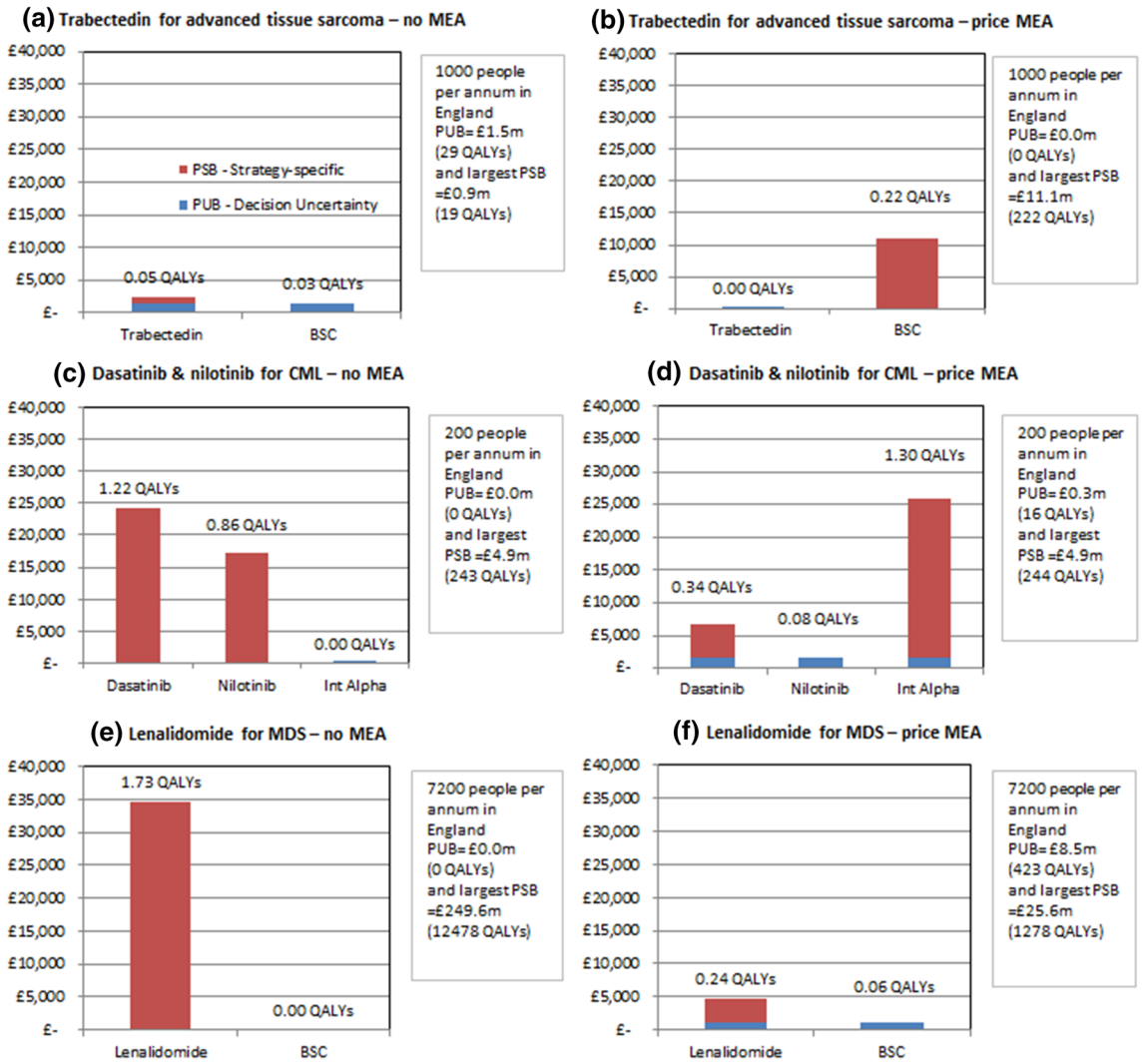

Fig. 2 Health technology assessment risk illustrated in three case studies. *BSC* best supportive care, *CML* chronic myeloid leukaemia, *MDS* myelodysplastic syndrome, *MEA* managed entry agreement, *PSB* payer strategy burden, *PUB* payer uncertainty burden, *QALY* quality-adjusted life-year

When decision makers are faced with highly uncertain decisions, they may be inclined to consider recommending technologies expected to be cost ineffective (i.e. with a positive PSB), under the condition that further research is undertaken. For example, the Cancer Drugs Fund currently funds drugs with large remaining uncertainty and that have plausible potential to be cost effective (and therefore can have a positive PSB), but under the condition that further data are collected [24]. The HTA risk analysis chart presented here provides greater transparency in presenting the implications of such a scheme through the PSB, which provides a window to deviations from risk neutrality. Say a technology has an ICER of £32,000 and the threshold is £30,000—this tells us nothing about the risk taken by recommending this technology. Our framework shows that the increase by recommending this technology in such a setting is quantified by the PSB.

The risk burden to the healthcare payer can be especially large where the affected patient population is sizeable. It is noteworthy that the appearance of the size of the stacked P-SUB bars is contingent on the scale used for the y-axis (this was particularly obvious when comparing the lenalidomide and trabectedin studies) and that the population value of the P-SUB should therefore be taken into account as a means of comparing P-SUBs across different HTAs. If well designed and used appropriately, MEA schemes can help to improve recommendations regarding new and existing technologies in a predictable, transparent and rational manner. While most HTA decisions are associated with some decision uncertainty, in our three examples there was little or no decision uncertainty and

consequently no PUB prior to the MEA because in these cases the original list prices were high in relation to the QALYs gained as estimated by the model and using the respective thresholds. Financial MEA schemes are likely to have an impact on the PUB, as shown in our case studies, highlighting the potential value in considering MEA schemes that have both financial and research elements.

A strength of this work is that the approach builds upon commonly used methods in technology appraisals and is therefore straightforward to use and understand. The P-SUB can be calculated directly and simply from a standard PSA, using value of information and expected net benefit analysis, enabling the analysis to be used routinely. Provided the PSA appropriately reflects uncertainty, then the resultant MEA analysis will also do so. However, if a model is believed by a committee to make implausible assumptions, or where important comparators are absent from the analysis, then the results derived by naïvely applying the HTA risk analysis chart framework would be misleading. In such cases, additional work is necessary to address the 'structural uncertainty' in the decision model and this is of course an area of active ongoing research [25–28]. For our framework to be of use in the context of substantial structural uncertainty, it is necessary to develop a PSA output that reflects that uncertainty. For example, in a context where two different options exist for a particular form or assumption (e.g. a survival curve functional form), then one approach would be to produce a set of PSA runs for each structure and form a judgement about the weighting of each set (perhaps based on statistical or expert considerations about plausibility), to generate a single PSA reflecting the structural uncertainty, a process called model averaging [27]. To the knowledge of the authors, this approach has not yet been used frequently within NICE decision making.

We wish to caution the reader from using our approach as a standalone method to assess the need for and value of MEAs. This paper is best viewed in the context of existing work on categorising MEA schemes [3, 4] and that describes the conditions under which research schemes may be of value [4, 5, 9, 12]. Claxton et al. [12] developed key principles and associated criteria that might guide 'only in research' and 'approval with research' recommendations. Garrison et al. [3] set out good practices in the assessment of desirability, design, implementation and evaluation of performance-based risk-sharing arrangements (PBRSAs). The authors provided definitions of different PBRSAs and a taxonomy of these, described their uses in different settings and jurisdictions and drew several conclusions, among them that research schemes should be assessed for their value of research and their quality and that ex post evaluations of any schemes should be set out at their design stage. Walker et al. [4] described a framework within which different coverage with evidence development decision options can be evaluated. The authors highlighted the importance of the technology's value, its associated value of research, the anticipated effect of coverage on further research as well as the costs of reversing, and the decision-makers ability to reverse, the decision. McKenna et al. [9] developed a checklist of assessment of different 'coverage with evidence development' schemes and applied it in case studies. The authors concluded that cost effectiveness is a necessary but not sufficient condition for a technology's recommendation for reimbursement. Rothery et al. [5] extended these research efforts to characterise the uncertainty and value of research issues specific to medical devices. When research schemes are considered, practical aspects such as the timing, the reversibility of the decision and ethical issues that may prevent research from being conducted once the technology is widely available become relevant. We have not provided a detailed review of guiding principles for the use of such schemes, as this can be found elsewhere. Further research should focus on the quantification of risk over the decision relevance horizon. It is also important to note that research schemes are only ever worthwhile if the decision maker is prepared to amend, including potentially reverse, a decision should the new evidence show that a funded technology is not, after all, cost effective. Such stopping rules need to be laid out at the design stage of MEAs [3]. The use and analysis of MEAs considers decision making as a series of related decisions rather than a one-shot choice.

These research efforts are particularly relevant in the currently changing environment in which a greater proportion of HTA submissions will be for technologies that have an immature evidence base, a trend that could lead to payers asking for more research in many situations, regardless of whether this is indicated or not. Routine consideration of the P-SUB, together with other decision algorithms that aid the assessment of when a research-based MEA scheme is indicated and, in fact, permissible, can aid the appropriate use of research-based schemes and thus prevent resources from being allocated to research projects that do not aid decision making. Further consideration is currently given to how the different frameworks could be integrated and implemented in the most efficient way. For example, manufacturers could be required to provide EVPI analysis (as is already the case in the Dutch HTA system [29]) and EVPPI analysis, with the rationale being that these analyses require only very little additional calculation; the design and valuation of research studies could be performed by a specialist group. We strongly advocate for further research into a comprehensive framework that unifies considerations about the assessment of the need for MEAs, the prerequisites for the use of MEAs, the design of suitable MEAs and the value of these MEAs using these existing algorithms as well as quantitative assessments.

Although our examples are based in the UK and with NICE technology appraisals, the framework is generalisable to any jurisdiction in which cost effectiveness (not necessarily QALY based) and uncertainty analysis are amongst the criteria for reimbursement decisions. To facilitate the adoption of this framework by reimbursement decision makers, further consideration would have to be given to jurisdiction-specific requirements, such as requirements for specific types of analysis (e.g. budget-impact analyses are required in some jurisdictions).

5 Conclusions

Decision makers are, rightly, often more cautious about claims made for new technologies. Rather than reject them, there may be circumstances where their introduction into practice can be warranted, provided the risk of doing so is transferred to those making the positive claims rather than public health systems. Our approach, the HTA risk analysis chart, helps decision makers identify those situations by presenting a standardised visualisation to show the need for and potential value of different classes of MEA schemes. Its use in HTA could ensure that MEAs are considered routinely, consistently and transparently, and it should prove particularly useful to both payers and manufacturers in the currently changing pharmaceutical environment.

Author Contributions Sabine Grimm drafted the main body of this manuscript, contributed to the development of methods and performed most of the analysis. Mark Strong and Alan Brennan drafted parts of this manuscript, provided feedback and contributed to the development of methods and the analysis. Allan Wailoo drafted parts of this manuscript, provided feedback and contributed to the development of methods.

Compliance with Ethical Standards

Data Availability Statement The data used for this study were made available by NICE and the companies involved in the described NICE technology appraisals. These data are not publicly available as they contain confidential information. More detailed description of the analysis can be found in the NICE Decision Support Unit report available at http://scharr.dept.shef.ac.uk/nicedsu/methods-development/managed-entry-agreements-mea/.

Funding Financial support for this study was provided by NICE through its Decision Support Unit. The views, and any errors or omissions, expressed in this article are those of the authors only. This report is independent research supported (in part) by the National Institute for Health Research (Mark Strong—Post-Doctoral Fellowship, PDF-2012-05-258). The views expressed in this publication are those of the author(s) and not necessarily those of the national health service, the National Institute for Health Research or the Department of Health. The funding agreements ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Publication of study results was not contingent on the sponsor's approval or censorship of the manuscript.

Conflict of interest Sabine Grimm, Mark Strong, Alan Brennan and Allan Wailoo have no conflicts of interest to declare.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. European Medicines Agency. Guideline on the scientific application and the practical arrangements necessary to implement Commission Regulation (EC) No 507/2006 on the conditional marketing authorisation for medicinal products for human use falling within the scope of Regulation (EC) No 726/2004. London: EMA; 2016.
2. European Medicines Agency. Adaptive Pathways 2014. http://www.ema.europa.eu/ema/index.jsp?curl=pages/regulation/general/general_content_000601.jsp. Accessed 21 Aug 2017.
3. Garrison LP Jr, Towse A, Briggs A, de Pouvourville G, Grueger J, Mohr PE, et al. Performance-based risk-sharing arrangements-good practices for design, implementation, and evaluation: report of the ISPOR good practices for performance-based risk-sharing arrangements task force. Value Health. 2013;16(5):703–19.
4. Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. Value Health. 2012;15(3):570–9.
5. Rothery C, Claxton K, Palmer S, Epstein D, Tarricone R, Sculpher M. Characterising uncertainty in the assessment of medical devices and determining future research needs. Health Econ. 2017;26(Suppl 1):109–23.
6. Grimm S, Strong M, Brennan A, Wailoo A. Framework for analysing risk in health technology assessments and its application to managed entry agreements. Sheffield: University of Sheffield; 2016.
7. Hollis A. Sustainable financing of innovative therapies: a review of approaches. Pharmacoeconomics. 2016;34(10):971–80.
8. Carlson JJ, Chen S, Garrison LP, Jr. Performance-based risk-sharing arrangements: an updated international review. Pharmacoeconomics. doi:10.1007/s40273-017-0535-z.
9. McKenna C, Soares M, Claxton K, Bojke L, Griffin S, Palmer S, et al. Unifying research and reimbursement decisions: case studies demonstrating the sequence of assessment and judgments required. Value Health. 2015;18(6):865–75.
10. Strong M, Oakley JE, Brennan A, Breeze P. Estimating the expected value of sample information using the probabilistic sensitivity analysis sample: a fast, nonparametric regression-based method. Med Decis Mak. 2015;35(5):570–83.
11. Eckermann S, Willan AR. Expected value of information and decision making in HTA. Health Econ. 2007;16(2):195–209.
12. Claxton K, Palmer S, Longworth L, Bojke L, Griffin S, McKenna C, et al. Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. Health Technol Asses. 2012;16(46):1–323.

13. Brennan A, Kharroubi SA. Efficient computation of partial expected value of sample information using Bayesian approximation. J Health Econ. 2007;26(1):122–48.
14. Raiffa H. Decision analysis: introductory lectures on choices under uncertainty. Reading: Addison-Wesley; 1968.
15. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
16. Wilson ECF. A practical guide to value of information analysis. Pharmacoeconomics. 2015;33(2):105–21.
17. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Med Decis Mak. 2014;34(3):311–26.
18. Strong M, Oakley JE. An efficient method for computing single-parameter partial expected value of perfect information. Med Decis Mak. 2013;33(6):755–66.
19. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Mak. 2004;24(2):207–27.
20. National Institute for Health and Care Excellence. Trabectedin for the treatment of advanced soft tissue sarcoma (TA185) 2010. London: NICE; 2017. https://www.nice.org.uk/Guidance/ta185. Accessed Mar 2017
21. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London: NICE; 2013. http://www.nice.org.uk/article/pmg9/chapter/1-Introduction. Accessed 21 Aug 2017.
22. National Institute for Health and Care Excellence. Dasatinib, high-dose imatinib and nilotinib for the treatment of imatinib-resistant chronic myeloid leukaemia (CML) (part review of NICE technology appraisal guidance 70), and dasatinib and nilotinib for people with CML for whom treatment with imatinib has failed because of intolerance (TA241) 2012. London: NICE; 2012. https://www.nice.org.uk/guidance/ta241. Accessed 21 Aug 2017.
23. National Institute for Health and Care Excellence. Lenalidomide for treating myelodysplastic syndromes associated with an isolated deletion 5q cytogenetic abnormality (TA322) 2014. London: NICE; 2014. https://www.nice.org.uk/guidance/ta322. Accessed 21 Aug 2017.
24. National Institute for Health and Care Excellence. Cancer Drugs Fund 2016. London: NICE; 2017. https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/nice-technology-appraisal-guidance/cancer-drugs-fund. Accessed Aug 2017.
25. Strong M, Oakley JE. When is a model good enough? Deriving the expected value of model improvement via specifying internal model discrepancies. SIAM/ASA J Uncertain Quantif. 2014;2(1):106–25.
26. Jackson CH, Bojke L, Thompson SG, Claxton K, Sharples LD. A framework for addressing structural uncertainty in decision models. Med Decis Mak. 2011;31(4):662–74.
27. Bojke L, Claxton K, Sculpher M, Palmer S. Characterizing structural uncertainty in decision analytic models: a review and application of methods. Value Health. 2009;12(5):739–49.
28. Strong M, Oakley JE, Chilcott J. Managing structural uncertainty in health economic decision models: a discrepancy approach. J R Stat Soc C-Appl. 2012;61:25–45.
29. Al M, Bindels J, Corro Ramos I, Feenstra T, Grutters J, Joore MA, et al. Onzekerheid en Value of Information Diemen: Zorginstituut Nederland; 2016. https://www.zorginstituutnederland.nl/publicaties/publicatie/2016/02/29/richtlijn-voor-het-uitvoeren-van-economische-evaluaties-in-de-gezondheidszorg. Accessed 21 Aug 2017.