RESEARCH ARTICLE

Health Economics WILEY

Development of a model to demonstrate the impact of National Institute of Health and Care Excellence cost-effectiveness assessment on health utility for targeted medicines

Daniel Gallacher[1] | **Nigel Stallard**[1] | **Peter Kimani**[1] |
Elvan Gökalp[2] | **Juergen Branke**[3]

[1]Warwick Medical School, University of Warwick, Coventry, UK

[2]School of Management, University of Bath, Bath, Somerset, UK

[3]Warwick Business School, University of Warwick, Coventry, UK

Correspondence
Daniel Gallacher,
Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK.
Email: d.gallacher@warwick.ac.uk

Abstract

Advances in medical technology have led to a better understanding of heterogeneity of diseases and patients, and to the development of targeted medicines. This development is beneficial to society but can come at an increased cost to pharmaceutical manufacturers due to the costs associated with developing and manufacturing a diagnostic test. For such medicines, the conventional pricing structure, where a therapy is approved if it is deemed cost-effective, may not appropriately incentivize targeted drug development. We model the decision-making processes for both the healthcare provider and the pharmaceutical manufacturer, capturing their main priorities, and populate it with information from a recent appraisal by the National Institute of Health and Care Excellence. Healthcare providers prefer a stratified drug to be developed for a subgroup of the population when the drug is on average effective in the subgroup but with a detrimental effect in the complement. Whilst pharmaceutical manufacturers' preferences are similar, regions of disagreement exist. We show how preferences can be aligned by either penalizing the development of a non-stratified drug or rewarding the development of a stratified drug. The cost and position of alignment depends on the true value of health to the healthcare provider, among other parameters.

KEYWORDS
cost-effectiveness analysis, health technology assessment, stratified therapies

1 | INTRODUCTION

Advancing medical knowledge and technologies are increasing understanding of patient heterogeneity. Patients previously thought to have the same disease and treated homogeneously can be discovered instead to have subtle differences and respond differently to different treatments (Bieber, 2013; Padmanabhan, 2014). An effective treatment for one patient

may not be as effective for others. To provide equal and optimal care for all, it is important to identify differences within disease groups and develop therapies to provide effective care for all patients. Heterogeneity may be identified through discovery of biomarkers including genetic testing, blood testing or medical imaging. Stratified therapies can be developed to specifically target biomarker positive patients (Beckman et al., 2011).

Stratified therapies may come at a higher cost to pharmaceutical manufacturers, with both the initial discovery of disease biomarkers and their routine detection incurring additional costs (ABPI, 2014). This may not be fully acknowledged in decision-making processes of healthcare providers. The National Institute of Health and Care Excellence (NICE) in England and Wales, currently gives no additional consideration to stratified treatments over conventional unstratified (full population) therapies and the economic evaluation of stratified medicines can bring additional challenges (Coyle et al., 2020; Hawkins & Scott, 2011; Shabaruddin et al., 2015). Pharmaceutical manufacturers may therefore be cautious to invest additional costs required to develop a stratified therapy without any additional expected reimbursement. Consequently, current practice may be stifling development of stratified therapies, with patients and healthcare providers losing out.

Financially, a pharmaceutical manufacturer may prefer to develop drugs for heterogeneous populations. Existing motivations for developing a stratified therapy may be that a biomarker subgroup is already established, or a therapy has failed to gain approval in a broader patient population. For example, CRYSTAL study data showed that combination therapy was only cost-effective for specific subgroups of patients in the original trial (Harty et al., 2018). However, a healthcare provider may need to provide greater incentive for continued and focused development of stratified therapy by pharmaceutical manufacturers, such as flexible pricing (Anonymous, 2013).

A number of authors have suggested modeling of the decision-making processes for stratified therapies. Sahlin and Hemerén present a decision theoretic model based on the idea of personalized medicine and discuss potential moral issues that may arise (Sahlin & Hermerén, 2012). Bardey and De Donder model the effect on genetic screening to identify patients for prevention methods, considering the perspective of the insurer (Bardey & De Donder, 2013).

Antoñanzas et al. model the decision of the health authority when faced with the decision whether to use a test to match patients to a treatment (Antoñanzas et al., 2015). They consider two treatments where each is most effective for a different subgroup of patients and conclude personalized medicine may impact drug development and reimbursement decisions. Zaric (2016) explores the impact to the payer of implementing precision medicine via a companion diagnostic test across four scenarios where a drug and biomarker test have already been developed.

In this paper we develop a model for the decision-making process for stratified therapies. This enables us to establish when a pharmaceutical manufacturer and a national decision-maker prefer either a stratified or unstratified therapy. Our model is distinct from the existing literature in that it focuses on the decision of NICE but also considers the view of the pharmaceutical manufacturer in the development of stratified therapies. In this work we show that as a consequence of the current processes of health technology assessment, preferences of the healthcare provider and pharmaceutical manufacturer for stratification for a particular therapy can be misaligned, and consider the conditions under which this misalignment occurs. We then explore solutions that reduce/remove this misalignment, increasing the motivation for developing stratified therapies and improving health care.

2 | UTILITY MODELS

To consider the impact of different methods of incentivization, we model the pharmaceutical manufacturer decision making process, and the preferences of the healthcare provider. Using decision theoretic methods (DeGroot, 2005; Minton et al., 1962; Oliehoek & Visser, 2010), we incorporate utility functions capturing the main factors considered by either a pharmaceutical manufacturer or healthcare provider in deciding to develop or reimburse a therapy.

The healthcare provider is motivated to provide the best healthcare or obtain the most quality adjusted life-years (QALYs) subject to its budget. However, when appraising a single therapy, NICE does not consider the cost implications on a micro-economic scale and does not undertake a cost minimization exercise. Instead, the selection of the most cost-effective therapies is indirectly achieved using willingness-to-pay thresholds, with decisions to fund a particular therapy being made on a case-by-case basis as new therapies become available. Meanwhile pharmaceutical manufacturers bear the upfront costs of developing a therapy, plus the potential cost of developing a biomarker test for a stratified therapy. Reimbursement for any drug developed must therefore cover the pharmaceutical manufacturer's initial investment, through a combination of large effect size leading to high price and/or large patient population.

Considering these two perspectives led us to develop the utility model described in the next two subsections. Model parameters are given in Table 1.

2.1 | Healthcare provider utility functions

The utility to the healthcare provider captures the value of the benefit to the population associated with the availability of the new drug less the cost that the healthcare provider must pay for the drug, both relative to existing care. Our theoretical "healthcare provider" represents both NICE, who are the decision maker in England and Wales, and the NHS, who deliver healthcare.

Let n denote the number of patients expected to receive the therapy across the lifetime of the treatment if approved for the full population. We define b to be the prevalence of a subgroup, with $1 - b$ the prevalence of its complement, where the treatment effect may vary across the subgroup and complement. We assume all patients fall into one of these two groups.

Let θ^S and θ^C denote the true average benefit (in QALYs) per patient in the subgroup and complement respectively. We assume this benefit is known, following estimation in a successful clinical trial assessing the efficacy of the treatment. For a stratified treatment, that is, one developed for treatment of the subgroup alone, the total health benefit (in QALYs) will then be $nb\theta^S$. For a non-stratified treatment, that is one developed in the full population, the total health benefit will be $n\theta^F$ where $\theta^F = b\theta^S + (1 - b)\theta^C$. If the true monetary value of a QALY is equal to u, with units £/QALY, then the value of the new stratified or non-stratified drug is respectively $nbu\theta^S$ or $nu\theta^F$. Treatment related inputs could be based on absolute value or relative to existing care.

Assume that the healthcare provider is willing to pay a total of k_F per QALY for an unstratified drug and its associated costs, and similarly k_S for a stratified drug. For our base scenario, we assume these to have the same value, however we model these separately to allow exploration of varying their values independently. To estimate the utility to the healthcare provider, we subtract these costs from the value of the benefit received by the healthcare provider for the therapy based on their willingness to pay threshold. This gives the following total utility for the full population:

TABLE 1 Definitions and values of parameters used in this study

Definition	Notation	Value used	Source
Population parameters			
Population size	n	1000	See text
Biomarker prevalence in population	b	0 to 1	
Actual effect in biomarker positive (subgroup)	θ^S	0.85 (QALYs)	Taken from NICE TA519
Actual effect in biomarker negative (complement)	θ^C	−1 to 1 (QALYs)	
Production costs			
Therapy production cost (per patient)	c_p	100 (£)	Estimate
Biomarker test production cost (per patient)	c_t	10 (£)	Estimate
Therapy development cost (total)	d	5,000,000 (£)	Estimate
Biomarker test development cost (total)	d_t	500,000 (£)	Estimate
Additional costs associated with treatment			
Incremental difference for one-off costs (per patient, e.g. administration, AE management)	c_a	−711 (£)	Taken from NICE TA519
Incremental costs per additional QALY gained (per patient, e.g. disease monitoring)	c_q	11,450.6 (£/QALY)	Taken from NICE TA519
Healthcare provider willingness to pay			
Threshold healthcare provider is willing to pay for stratified treatment, per QALY gained	k_S	48,000 (£/QALY)	Taken from NICE TA519
Threshold healthcare provider is willing to pay for a non-stratified treatment, per QALY gained	k_F	48,000 (£/QALY)	Taken from NICE TA519
True value of QALY to healthcare provider	u	60,000 (£)	Estimate

Abbreviations: AE: adverse event, NICE, National Institute of Health and Care Excellence; QALY: quality adjusted life-year.

$$U^{HP,F} = n\left(u - k_F\right)\theta^F. \tag{1}$$

The utility function for the healthcare provider for a stratified therapy has a form similar to the utility for the full population, but contains subgroup-specific terms, and is given by

$$U^{HP,S} = n\,b\left(u - k_S\right)\theta^S. \tag{2}$$

It would not make sense for $u < k_F$ since it is illogical for the healthcare provider to be willing to pay more than it values the health gain. A healthcare provider has a limited budget meaning it cannot afford to pay u for every treatment as it must ensure it can consider future treatments. Furthermore, if $u = k_F$ the healthcare provider has no preference whether new treatments are developed or reimbursed, nor any incentive to consider new therapies, since it receives the same value for money regardless. Hence $u > k_F$ so that the healthcare provider has an incentive to adopt new and better therapies. Similarly $u > k_S$.

2.2 | Pharmaceutical manufacturer utility functions

For the pharmaceutical manufacturer, we assume a utility function equal to the total profit from treating the full population or identified subgroup with the drug. This is equal to the average price paid by the healthcare provider to the pharmaceutical manufacturer for the therapy minus manufacturing cost (c_p), multiplied by the number of patients, n, less the development costs, d.

Let v_S and v_F denote the prices of stratified and conventional therapies respectively accounting for the total that the healthcare provider is willing to pay for the overall benefit with c_q denoting the average additional cost incurred per patient for each incremental QALY, and c_a the average additional cost incurred, independent of QALY benefit. We assume the pharmaceutical manufacturer negotiates the maximum price the healthcare provider is willing to pay. Thus v_S and v_F are given by

$$v_F = \left(k_F - c_q\right)\theta^F - c_a \tag{3}$$

$$v_S = \left(k_S - c_q\right)\theta^S - c_a. \tag{4}$$

We assume that the diagnostic and drug are developed and sold by the same pharmaceutical manufacturer. This means that we do not need to assume the healthcare provider pays separately for the diagnostic test.

For a non-stratified medicine developed for the whole patient population, the utility to the pharmaceutical manufacturer is thus

$$U^{PM,F} = n\left(v_F - c_p\right) - d. \tag{5}$$

The utility for a stratified therapy is similar, but includes the costs of developing (d_t) and producing (c_t) the biomarker test:

$$U^{PM,S} = n\left(b\left(v_S - c_p\right) - c_t\right) - d - d_t \tag{6}$$

3 | RESULTING PREFERENCES FOR PHARMACEUTICAL MANUFACTURER (PM) AND HEALTHCARE PROVIDER (HP)

3.1 | PM preference general form

The pharmaceutical manufacturer will prefer to develop a stratified therapy whenever

$$U^{PM,S} > U^{PM,F}. \tag{7}$$

From Equations (5) and (6), that is whenever

$$\theta^C < \frac{c_a + c_p - \dfrac{c_t + \dfrac{d_t}{n}}{1-b} - \dfrac{b\left(k_F - k_S\right)\theta^S}{1-b}}{k_F - c_q}. \tag{8}$$

If $k_F = k_S$, this simplifies to

$$\theta^C < \frac{c_a + c_p - \dfrac{c_t + \dfrac{d_t}{n}}{1-b}}{k_F - c_q}. \tag{9}$$

In order for the pharmaceutical manufacturer to be motivated to develop the therapy, at least one of $U^{PM,F}$ and $U^{PM,S}$ must be > 0.

3.2 | HP preference general form

Similarly, assuming $u > k_F$, the healthcare provider will prefer to have a stratified therapy whenever

$$\theta^C < \frac{b\,\theta^S\left(k_F - k_S\right)}{(1-b)\left(u - k_F\right)}. \tag{10}$$

If $k_F = k_S$ this simplifies to $\theta^C < 0$.

3.3 | Alignment of pharmaceutical manufacturer and healthcare provider preferences

From Equations (8) and (10) the preferences of the healthcare provider and pharmaceutical manufacturer will align when

$$\frac{b\,\theta^S\left(k_F - k_S\right)}{(1-b)\left(u - k_F\right)} = \frac{c_a + c_p - \dfrac{c_t + \dfrac{d_t}{n}}{1-b} - \dfrac{b\left(k_F - k_S\right)\theta^S}{1-b}}{k_F - c_q}. \tag{11}$$

Setting $k_S = k_F$ means the left side of the equation reduces to zero, and the right simplifies so that

$$\frac{c_t + \dfrac{d_t}{n}}{1-b} = c_a + c_p. \tag{12}$$

This would be true when the costs of treating a patient in the complement, $(1-b)(c_a + c_p)$, are equal to the cost per patient in the full population of developing and producing the biomarker test, $c_t + d_t/n$.

As the decision to develop the therapy either for a biomarker positive population or for a wider population lies with the pharmaceutical manufacturer, the healthcare provider may be left with a suboptimal outcome. This is explored in detail in the example below.

3.4 | Retrospective example: pembrolizumab for advanced urothelial carcinoma

To assess the characteristics of our model with realistic values, as our base scenario, we retrospectively use the setting and parameter values of the publicly available information from the NICE single technology appraisal of pembrolizumab for

treating locally advanced or metastatic urothelial carcinoma after platinum containing chemotherapy (Anonymous, 2017; Bellmunt et al., 2017; Gallacher et al., 2019). These are shown in Table 1, alongside details explaining their source. Later, these parameters are varied in sensitivity analyses.

In this appraisal there was no restriction of pembrolizumab to specific subgroups of patients. However, the same therapy is restricted to patients with a specific biomarker (PD-L1 status) in other indications. PD-L1 status is assessed as the combined positive score (CPS), measuring the number of PD-L1 positive cells relative to the total number of tumor cells. Clinical outcomes from the KEYNOTE-045 trial of pembrolizumab (Bellmunt et al., 2017) for the PD-L1 subgroups were presented for patient groups with ≥1% CPS and ≥10% CPS, with respective prevalence (b) of approximately 40% and 30% from the whole patient population. However, QALY estimates from TA519 were only available for the full trial population.

The mechanism of action of the therapy is consistent with these subgroups, and the KEYNOTE-045 trial of pembrolizumab protocol specified that it would explore these subgroups. The hazard ratio for overall survival did show a trend with PD-L1 (Bellmunt et al., 2017). However, the researchers did not find a statistically significant interaction effect between PD-L1 status and pembrolizumab in KEYNOTE-045, and chose to seek approval for the wider population, ignoring PD-L1 status (Anonymous, 2017).

This case study was selected because of the availability of the parameter values necessary for our model. We use this case study to illustrate the impact of a range of different sizes of effect in the PD-L1 negative subgroup, varying the value of θ^C, including when there is equal efficacy to the subgroup, generalizing beyond the pembrolizumab example. Negative efficacy represents the potential negative effects of pembrolizumab (low absolute efficacy and adverse events) relative to existing care. We do not mean to suggest that pembrolizumab should have been approved as a stratified therapy for this indication.

The population size, n, is the number of patients likely to receive the therapy across the lifetime of the therapy. The company predicted approximately 500 patients would be eligible for therapy annually. Given the existence of approved similar therapies and evolving treatment pathway, we set n as 1000.

We used the value of 48,000 as the threshold (k_S, k_F) as this matches the incremental cost-effectiveness ratio (ICER) from the company base case analysis in their initial submission, and assumes the pharmaceutical manufacturer will allow for some uncertainty in their modeling, rather than hitting the £50,000 per incremental QALY threshold for end of life therapies.

For c_a we added the incremental costs that all patients would incur regardless of the level of benefit received (terminal care cost, post-discontinuation cost, adverse event cost). For c_q we combined the incremental costs that were affected by QALY benefit (disease management cost, drug administration cost), and divided by the total number of incremental QALYs. For other applications of our model, it may make sense to make administration cost independent of benefit.

We set u as £60,000 initially, implying the healthcare provider makes meaningful gains of approximately 20% per QALY compared to the default willingness-to-pay threshold, and consider alternative values in sensitivity analyses reported below.

3.5 | Results of example

Figure 1 shows the range of values of b and θ^C where there is preference for stratification for each of the healthcare provider and pharmaceutical manufacturer. The healthcare provider boundary of preference is shown by the purple line, horizontal at the line of no effect in the complement group ($\theta^C = 0$). When there is a positive treatment effect in the complement, the healthcare provider would prefer to give these patients access to the treatment, that is for the drug to be developed for the full population, and when there is a negative effect, the healthcare provider would prefer stratification.

The boundary of preference for the pharmaceutical manufacturer is shown by the pink curve. Above the curve, the company prefers not to stratify, whilst below it prefers to stratify. The region indicated with white lines portrays the values of b and θ^C where it is not in the pharmaceutical manufacturer's interest to develop a therapy (either stratified or for the full population) as they are unable to recoup development costs and so may not develop a drug at all.

Details underlying Figure 1 are given in Supporting Information.

The disagreement between the pharmaceutical manufacturer and healthcare provider is shown by the region between the purple and pink curves in Figure 1. The different colors in this region indicate the magnitude of the loss to the healthcare provider when the pharmaceutical manufacturer chooses to develop the therapy in line with their own preferences. The region is characterized by a weak negative effect in the complement population. The region expands
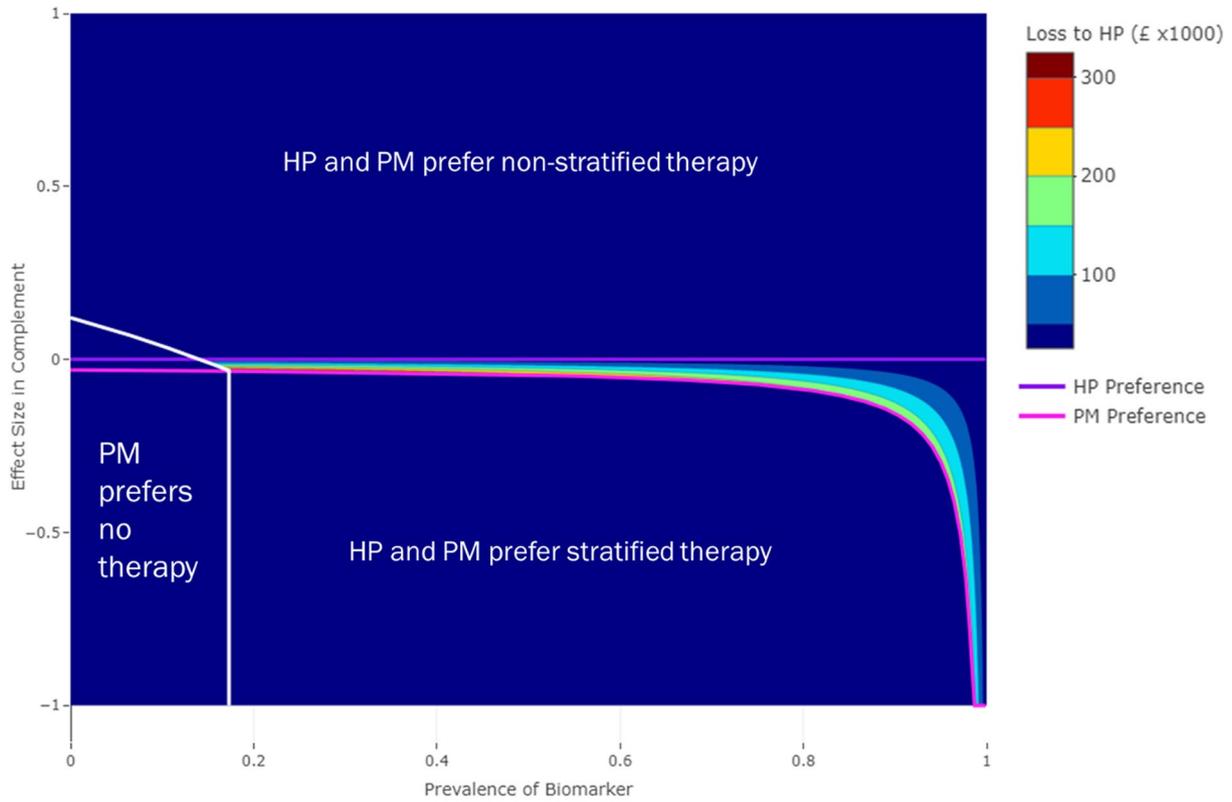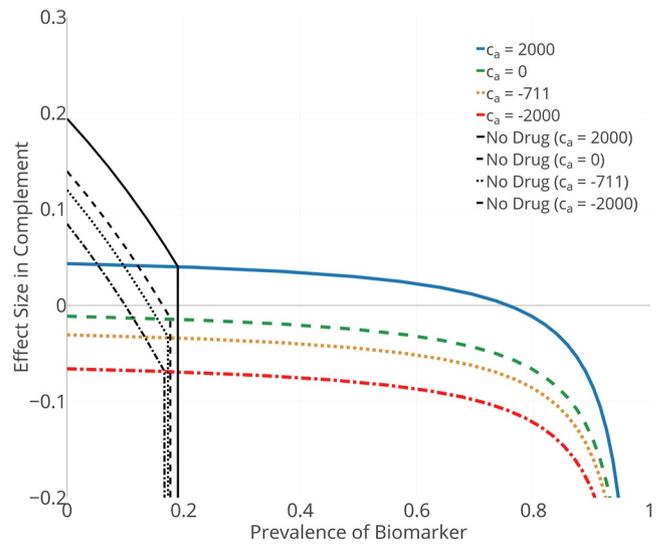
FIGURE 1 Contour plot of healthcare provider (HP) and pharmaceutical manufacturer (PM) preferences, and healthcare provider losses if pharmaceutical manufacturer and healthcare provider preferences do not align

FIGURE 2 Demonstrating preferences for the pharmaceutical manufacturer for different values of c_a

to include much stronger values of negative effect when the prevalence of the biomarker negative population is much smaller relative to the biomarker positive population.

Figure 2 shows the sensitivity of the pharmaceutical manufacturer initial preference to the parameter c_a with higher positive values leading the pharmaceutical manufacturer to prefer stratification even when there is a small benefit of the treatment in the complement population.

Additional sensitivity analyses to both the base case preferences are shown in the appendix.

4 | METHODS FOR ALIGNING PREFERENCES OF HEALTHCARE PROVIDER AND PHARMACEUTICAL MANUFACTURER

We considered three approaches to aligning the preferences of the healthcare provider and pharmaceutical manufacturer.

4.1 | Increasing the price of stratified therapies

Our first approach is to increase the amount the healthcare provider is willing to pay for a stratified therapy (k_S).

Rearranging Equation (8), the healthcare provider's and pharmaceutical manufacturer's preferences align if

$$k_S = k_F - \frac{(1-b)(u - k_F)\left(c_a + c_p - \frac{c_t + \frac{d_t}{n}}{1-b}\right)}{b\theta^S (u - c_q)}. \tag{13}$$

Returning to our pembrolizumab example, Figure 3 shows the changes to preferences of both the healthcare provider and pharmaceutical manufacturer when k_S is given by Equation (13) for a range of values of b. Resulting values of k_S are given in Table 2. The preferences now align, with the points at which they change shown by the solid pink line. This alignment has come at a cost to the healthcare provider who has had to compromise on their initial preference. The region indicated by the white line, where the pharmaceutical manufacturer would prefer not to develop the drug, has slightly reduced.
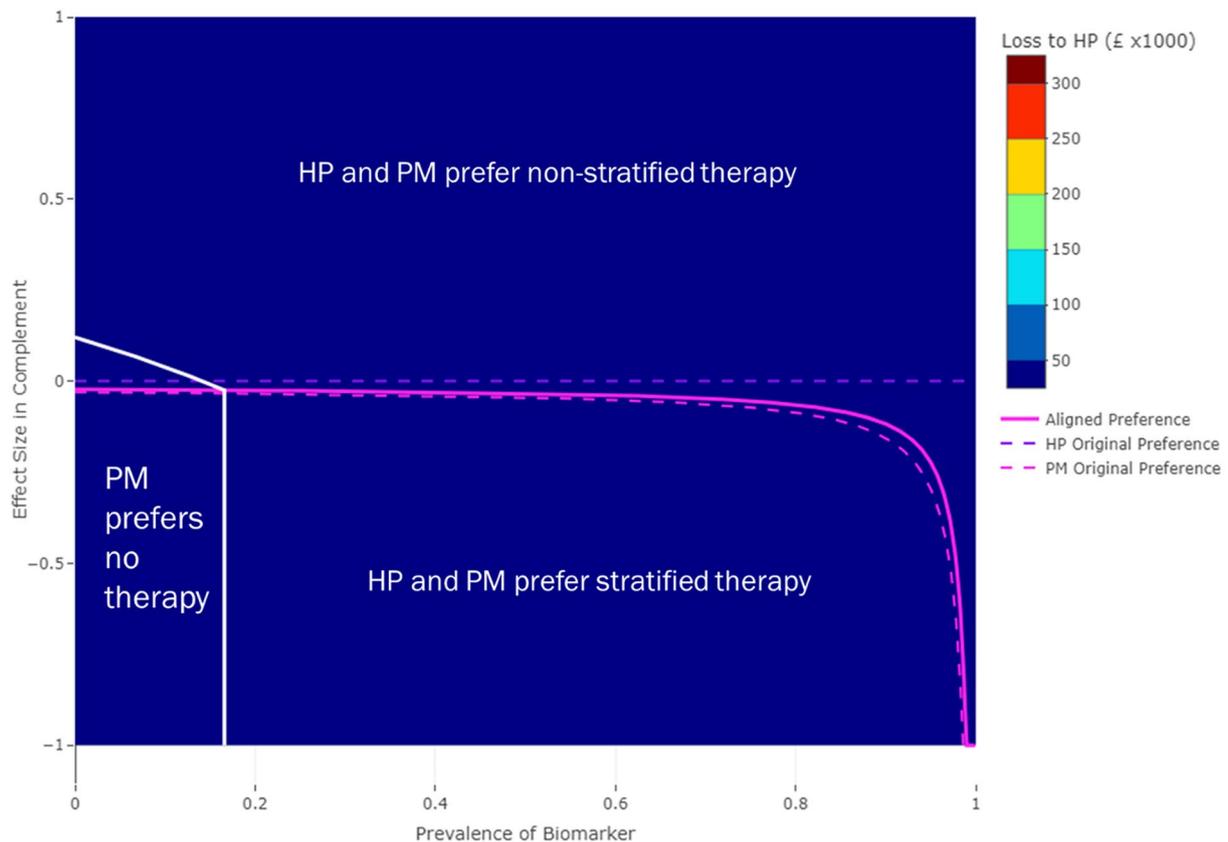

FIGURE 3 Alignment of solutions 1 and 2. HP, healthcare provider; PM, pharmaceutical manufacturer

TABLE 2 Resulting values for solutions 1 and 2 to align preferences for varying prevalence across the base case

b	0.05	0.1	0.2	0.3	0.4	0.5
Solution 1 parameter values (£/QALY)						
k_F	48,000	48,000	48,000	48,000	48,000	48,000
k_S	54,342 (+13.2)	51,082 (+6.4%)	49,452 (+3.0%)	48,909 (+1.9%)	48,637 (+1.3%)	48,474 (+1.0%)
Solution 2 parameter values (£)						
d_t	500,000	500,000	500,000	500,000	500,000	500,000
$d_{t,HP}$	269,528 (54%)	261,976 (52%)	246,874 (49%)	231,772 (46%)	216,670 (43%)	201,568 (40%)
$d_{t,PM}$	230,472 (46%)	238,024 (48%)	253,126 (51%)	268,228 (54%)	283,329 (57%)	298,432 (60%)

Abbreviation: QALY: quality adjusted life-year.

4.2 | Contributing to biomarker development costs

Our second approach is for the healthcare provider to make an upfront contribution toward biomarker test development costs. Let $d_{t,HP}$ and $d_{t,PM}$ denote contributions to the biomarker test development costs made by the healthcare provide and pharmaceutical manufacturer respectively, with $d_{t,HP} + d_{t,PM} = d_t$. We later discuss how this is different to the earlier solution.

Adapting Equation (11) to incorporate these contributions, the preferences will align when

$$d_{t,PM} = \frac{n}{u - c_q}\left(\frac{d_t\left(k_F - c_q\right)}{n} + \left(u - k_F\right)\left(b\theta^S\left(k_S - k_F\right) + (1-b)\left(c_a + c_p\right) - c_t - b\theta^S\left(k_F - k_S\right)\left(k_F - c_q\right)\right) \right). \quad (14)$$

With this value of $d_{t,PM}$ and $d_{t,HP} = d_t - d_{t,PM}$, we obtain alignment identical to Solution 1.

Table 2 shows resulting values of $d_{t,HP}$ and $d_{t,HP}$ for certain values of b. The healthcare provider contribution is largest when b is small, since the pharmaceutical manufacturer has a small patient pool to recover costs from and the drug price is independent of the population size.

4.3 | Penalizing negative effects

In our pembrolizumab example the trend of increasing efficacy with PD-L1 positivity is accompanied by a positive effect in PD-L1 negative patients. In other settings, the effect in biomarker negative patients may be negative. In this case desire for a stratified medicine is particularly great. Our third approach encourages the pharmaceutical manufacturer to develop a stratified therapy to reduce the number of patients experiencing negative effects. We add a penalty term P to the pharmaceutical manufacturer utility function for the full population, which magnifies the effect of any negative experiences. For simplicity, we assume that the negative effect only occurs in the complement population.

Suppose that the price for a non-stratified drug is now

$$v_{F,P} = \left(k_F - c_q\right)\theta^F - c_a + P(1-b)\theta^{C-} \quad (15)$$

where θ^{C-} is min $\left\{0, \theta^C\right\}$, thus there is no penalty if θ^C is positive.

The form of the pharmaceutical manufacturer utility functions remain as per Equations (5) and (6), as the change is captured through v_F, which influences $U^{PM,F}$. Meanwhile, for the healthcare provider, $nP(1-b)\theta^{C-}$ is subtracted from the previous definition of $U^{HP,F}$.

Setting an optimal value for P was not straightforward. We set $P = u - k_F$ since setting P to a value greater than $u - k_F$ leads to the healthcare provider preferring a treatment that performs less well in the complement population because of the large penalty. In order to achieve its aim, the value of P must be sufficiently large, and so we maximized its value, whilst still dependent on u. This also avoids selecting an arbitrary value for P. We explore other definitions and values of P in Appendix C in Supporting Information.

The resulting plot is almost identical to Figure 2 and is shown in Supporting Information Figure C2. The line for the healthcare provider should be interpreted differently to the previous figures. Whereas before it indicated the difference

between a region of preference for stratification and a region for no stratification, here it indicates the difference between a region of preference for no stratification (above) and a region where the healthcare provider has no preference (below). The alignment is identical to the first two solutions, however the region of no development has grown slightly.

Figure 4 demonstrates how the alignment of preference varies for different values of u, with higher values bringing the place of alignment closer to the original preference of the healthcare provider. Table 3 shows how the costs associated with the solutions increase as u increases.

5 | DISCUSSION

We have modeled utilities representing the values of both the healthcare provider and the pharmaceutical manufacturer. Each stakeholder's preference for a stratified or unstratified therapy was calculated using parameter values from a recent NICE single technology appraisal, whilst the subgroup:complement prevalence and complement treatment efficacy both varied. We demonstrated that for certain combinations of subgroup prevalence and complement efficacy, there is misalignment of interests between the healthcare provider and pharmaceutical manufacturer. This disagreement could lead the pharmaceutical manufacturer to develop a treatment as an unstratified therapy whilst the healthcare provider would rather it be developed as a stratified therapy, meaning the healthcare provider is not receiving its preferred outcome.

We have then explored potential ways of aligning the preferences of the two stakeholders and demonstrated the financial impact of each. We have proposed three possible modifications of the reward structure to bring these preferences into alignment. All three solutions achieve an identical compromise in the positioning of the line of alignment when the region of disagreement falls entirely within the area of $\theta^C < 0$, for a particular value of u. Our research builds on existing literature which has considered the challenges surrounding stratified therapies (ABPI, 2014; Wilsdon et al., 2018) and recommends collaboration between stakeholders (Attar et al., 2019; Cope et al., 2018).

Both increasing the willingness-to-pay threshold, or contributing in terms of a lump sum, see the healthcare provider preference line shift to compromise with pharmaceutical manufacturer, and "meet in the middle". The proximity of the compromise to either of the original healthcare provider or pharmaceutical manufacturers preference depends on u. For large u the compromise is very close to the healthcare providers original preference, however this comes at a very high financial cost. Whilst such an approach may prove necessary, it potentially leaves the process open to manipulation with a fair compromise relying on honest and transparent reporting of development costs including details on how these are shared across different markets/healthcare providers. It appears that the healthcare provider is now happy to unethically prefer cases where patients experience harm. However, the compromise is an improvement on current practice, reducing the frequency of occasions where the healthcare provider is presented with an unstratified therapy when they would prefer a stratified therapy, and negative effects are experienced by patients. If the healthcare provider could afford to spend even more, then they could incentivize the pharmaceutical manufacturer to align to the original preference of the healthcare provider.
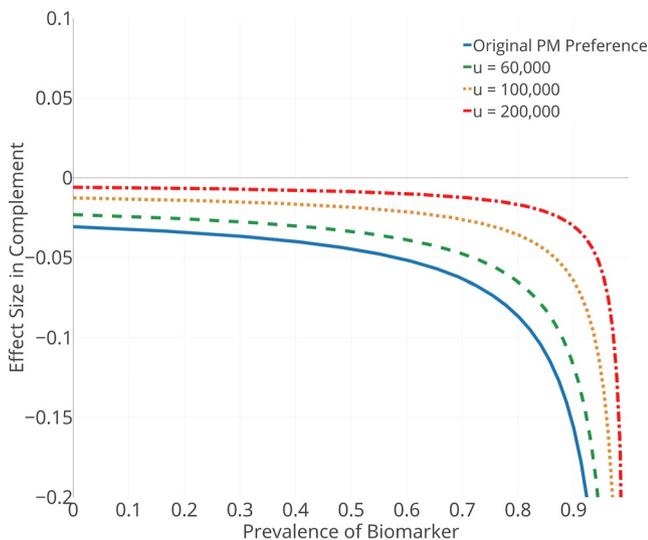

FIGURE 4 Position of alignment of preference for different values of u, the true value of a QALY to the healthcare provider. PM, pharmaceutical manufacturer

TABLE 3 Resulting values for three solutions for different values of u and b

Value of QALY to healthcare provider	Prevalence of subgroup	Sol 1: Raising threshold for stratified therapy	Sol 2: Upfront contribution	Sol 3: Penalty term
u	b	$k_S\left(k_F = 48{,}000\right)$	$d_{t,HP}\left(d_t = 500{,}000\right)$	$p(1-b)$
60,000	0.05	54,342	269,528	11,400
60,000	0.1	51,082	261,976	10,800
60,000	0.25	49,126	239,323	9000
60,000	0.4	48,637	216,670	7200
100,000	0.05	63,067	640,359	49,400
100,000	0.1	55,323	622,419	46,800
100,000	0.25	50,676	568,598	39,000
100,000	0.4	49,514	514,777	31,200
200,000	0.05	68,684	879,071	144,400
200,000	0.1	58,052	854,443	136,800
200,000	0.25	51,673	780,559	114,000
200,000	0.4	50,078	706,675	91,200

Abbreviation: QALY: quality adjusted life-year.

As far as we are aware, the approach of contributing to biomarker test development costs would set a new precedent in how pharmaceutical manufacturers received reimbursement from the NHS/NICE. However, there is precedent for NICE accepting a higher price for some types of targeted therapies as it has a separate appraisal route for highly specialized technologies (HST), which are identified on criteria including: small and distinct population, chronic and disabling disease, and an unmet need. HSTs are not subject to the same willingness to pay threshold as routine appraisals and use a higher threshold of £100,000 per QALY gained. Occasionally therapies have extended negotiation periods due to a conflict of valuations between NICE and the manufacturer, such as Orkambi for cystic fibrosis, which presumably end in a compromise in excess of the standard cost-effectiveness thresholds. It is possible that paying more per QALY for stratified therapies which affect smaller numbers of people may be a natural extension of NICE's current practices, and should be assessed on a more continuous scale than current dichotomization of the HST and single technology appraisal willingness-to-pay thresholds.

A limitation of introducing a penalty term for negative effects is that it is difficult to identify which patients experience a negative effect. It is unlikely a pharmaceutical manufacturer would seek approval for a subgroup who experience negative effects, nor that a healthcare provider would approve. However, since effects are measured at population level, a subgroup who experience negative effects could go undetected amongst a diluted population net benefit. A pharmaceutical manufacturer is unlikely to try too hard to identify this subgroup if they will then incur penalization. There are two simple ways that NICE appraisals could identify negative effects on a population level. First, common to all appraisals is the utility decrement for adverse events. These are often calculated and applied separately to other QALY parameters. However, the magnitude and influence of this form of penalization are insufficient to always outweigh the costs of stratification, such as the costs of developing a biomarker test, and may be ineffective at guiding the pharmaceutical manufacturer to the preference of the healthcare provider. The disutility applied for adverse events could easily be multiplied by a penalty term increasing its influence, which would capture some, and potentially all, of the negative effects experienced by patients. Secondly, certain modern treatments, including immunotherapies such as pembrolizumab, often appear to perform worse for patients in the first couple of months of follow-up. For immunotherapies, this is because of their mechanism of action relies on utilizing the body's immune system to kill cancer cells. As these treatments are novel, they are regularly compared to chemotherapies, which have a very instant and aggressive impact. This difference in mechanisms means that some patients on the immunotherapies experience progressive disease or death before the treatment has time to have effect. Hence there is a period of follow-up where the incremental QALYs for the novel therapy are negative compared to the reference treatment. These negative QALYs could also be penalized in appraisals where they occur.

Antoñanzas explores the impact of different kind of penalty, where pharmaceutical companies are penalized for each patient on whom their treatment fails (Antoñanzas et al., 2018). They report that this pay-for-performance style of penalization can be effective in encouraging stratified therapies, and that providing incentives to pharmaceutical companies will improve health outcomes. Implementing this approach has difficulties as assessing efficacy can be subjective when using outcomes such as disease progression and treatment response, and without knowledge of how a patient would otherwise have fared if given alternative treatment. It could also lead to pharmaceutical companies discouraging use of their treatments among the most critically ill, despite these patients potentially having the most to gain. For example, increasing the penalty for negative effects means the risk may outweigh the reward for unhealthy patients, discouraging the pharmaceutical company from treating these 'high risk' patients.

The difference between the upfront contribution and an amortized contribution is largely risk. If the estimates are all accurate, then they should be equivalent. However if, among other variation, the true number of patients is less than the estimate, the healthcare provider would be better paying the amortized amount. Alternatively, if the true number is higher than estimated, the healthcare provider achieves a better deal by paying their share of the development costs upfront. This uncertainty is increased due to the potential for new treatments to be developed.

One major difference between our penalty solution and the solutions based on the healthcare provider contributing toward the biomarker test development costs, whether amortized or upfront, is the impact on the size of the region of the values of b and θ^C for which the drug gets developed. Whilst the healthcare provider would ideally like to avoid the additional contribution, this risks a drug not being developed if the pharmaceutical manufacturer sees its potential rewards reduced.

The optimal solution for the healthcare provider may depend on the value of b, with a penalty preferred when the drug is likely to be developed anyway, and a contribution toward biomarker test development costs preferred when the drug is in danger of not being developed.

Our model is generalizable to other healthcare systems which appraise therapies on their cost-effectiveness on the QALY scale. It is a useful starting point to begin the discussion between healthcare providers and pharmaceutical manufacturers in cooperating toward the development of stratified therapies.

5.1 | Limitations

Our modeling requires specification of several unknown parameters and relies on plausible estimates or calculations to obtain values. Certain parameters, such as the value of the QALY to the healthcare provider, and costs of developing a therapy and the biomarker test, are influential parameters in the model and can vary hugely across diseases. Even if these costs were known, a pharmaceutical manufacturer might not disclose what proportions of these costs they expect to reclaim from specific countries.

The population size is also unknown, though pharmaceutical manufacturers are likely to have some expectation on the number of patients they expect to receive their product. Areas of uncertainty could include the presence of comparator therapies and development of future therapies.

Our healthcare provider utility depends on the true value of a QALY, u. Presumably NICE values a QALY at least at the value it is willing to pay, and potentially more than the established thresholds. The value of u does not alter the range of drugs for which the healthcare provider would prefer stratification, and simply scales the loss when the preference of the pharmaceutical manufacturer does not match that of the healthcare provider. In this sense the choice of u is arbitrary. Changing the value of a QALY to the healthcare provider, does however change the amount the healthcare provider is willing to pay or charge to incentivize the pharmaceutical manufacturer in order to bring their preferences into alignment. The larger u the larger the incentive, and correspondingly the closer the aligned preferences are to the original preference of the healthcare provider.

We have considered the impact of a drug in isolation, and assumed that the pharmaceutical manufacturer does not already have any stake in the existing therapy. We do not consider uncertainty in the parameters for drug efficacy, or on the potential outcome of clinical trials that are focused either on a subgroup or broader population. Our models could be extended to incorporate this uncertainty following the approach taken by Ondra et al. (2016).

Coyle et al. consider the possibility of leakage of stratified therapies, where treatments only deemed cost-effective for specific groups of patients are given by doctors to a wider population, which is not considered in this paper (Coyle et al., 2003).

6 | CONCLUSIONS

We have developed a model and illustrated it using a NICE single technology appraisal. We have shown that a misalignment of preferences for a stratified therapy can result in the healthcare provider losing health utility as it receives the therapy as developed by the pharmaceutical manufacturer. We implemented three solutions to align healthcare provider's and pharmaceutical manufacturer's preferences, which achieved identical alignment but had different effects on the pharmaceutical manufacturer's decision whether to develop the therapy at all. The positioning of resulting alignment depends on multiple parameters, including the true value of a QALY to the healthcare provider, which is difficult to quantify.

CONFLICT OF INTERESTS
All authors confirm they have no conflict of interests.

ETHICS STATEMENT
No ethics approval was necessary for this study.

AUTHOR CONTRIBUTION
Nigel Stallard, Juergen Branke, Peter Kimani and Elvan Gökalp generated the research idea. Juergen Branke and Elvan Gökalp completed background work. Daniel Gallacher implemented the model programming, identified model inputs and drafted the manuscript.

All authors contributed to all stages of the research and in producing the manuscript and approving the final version.

DATA AVAILABILITY STATEMENT
Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID
Daniel Gallacher https://orcid.org/0000-0003-0506-9384
Nigel Stallard https://orcid.org/0000-0001-7781-1512
Peter Kimani https://orcid.org/0000-0001-8200-3173
Elvan Gökalp https://orcid.org/0000-0001-9045-3860
Juergen Branke https://orcid.org/0000-0002-4343-5878

REFERENCES
ABPI. (2014). *Stratified medicine in the NHS: An assessment of the current landscape and implementation challenges for non-cancer applications*.
Anonymous (2013). *Realising the potential of stratified medicine*: Academy of Medical Sciences.
Anonymous (2017). *Committee Papers: Pembrolizumab for previously treated advanced or metastatic urothelial cancer*. Available from.https://www.nice.org.uk/guidance/ta519/documents/committee-papers
Antoñanzas, F., Juárez-Castelló, C. A., & Rodríguez-Ibeas, R. (2015). Some economics on personalized and predictive medicine. *The European Journal of Health Economics*, *16*(9), 985–994.
Antoñanzas, F., Rodríguez-Ibeas, R., & Juárez-Castelló, C. A. (2018). Personalized medicine and pay for performance: Should pharmaceutical firms be fully penalized when treatment fails? *PharmacoEconomics*, *36*(7), 733–743.
Attar, S., Poustie, V., Smye, S., Beety, J., Hawcutt, D., Littlewood, S., Oni, L., Pirmohamed, M., & Beresford, M. W. (2019). Working together to deliver stratified medicine research effectively. *British Medical Bulletin*, *129*(1), 107–116.
Bardey, D., & De Donder, P. (2013). Genetic testing with primary prevention and moral hazard. *Journal of Health Economics*, *32*(5), 768–779.
Beckman, R. A., Clark, J., & Chen, C. (2011). Integrating predictive biomarkers and classifiers into oncology clinical development programmes. *Nature Reviews Drug Discovery*, *10*(10), 735–748.
Bellmunt, J., De Wit, R., Vaughn, D. J., Fradet, Y., Lee, J.-L., Fong, L., Vogelzang, N. J., Climent, M. A., Petrylak, D. P., Choueiri, T. K., Necchi, A., Gerritsen, W., Gurney, H., Quinn, D. I., Culine, S., Sternberg, C. N., Mai, Y., Poehlein, C. H., Perini, R. F., & Bajorin, D. F. (2017). Pembrolizumab as second-line therapy for advanced urothelial carcinoma. *New England Journal of Medicine*, *376*(11), 1015–1026.

Bieber, T. (2013). *Stratified medicine: A new challenge for academia, industry, regulators and patients*. Future Medicine. https://doi.org/10.2217/9781780843186

Cope, A. P., Barnes, M. R., Belson, A., Binks, M., Brockbank, S., Bonachela-Capdevila, F., Bonachela-Capdevila, F., Carini, C., Fisher, B. A., Goodyear, C. S., Emery, P., Ehrenstein, M. R., Gozzard, N., Harris, R., Hollis, S., Keidel, S., Levesque, M., Lindholm, C., McDermott, M. F., Isaacs, J. D. (2018). The RA-MAP consortium: A working model for academia–industry collaboration. *Nature Reviews Rheumatology*, *14*(1), 53–60.

Coyle, D., Buxton, M. J., & O'Brien, B. J. (2003). Stratified cost-effectiveness analysis: A framework for establishing efficient limited use criteria. *Health Economics*, *12*(5), 421–427.

Coyle, D., Durand-Zaleski, I., Farrington, J., Garrison, L., Graf von der Schulenburg, J.-M., Greiner, W., Longworth, L., Meunier, A., Moutié, A.-S., Palmer, S., Pemberton-Whiteley, Z., Ratcliffe, M., Shen, J., Sproule, D., Zhao, K., & Shah, K. (2020). HTA methodology and value frameworks for evaluation and policy making for cell and gene therapies. *The European Journal of Health Economics*, *21*(9), 1421–1437.

DeGroot, M. H. (2005). *Optimal statistical decisions*. John Wiley & Sons.

Gallacher, D., Armoiry, X., Auguste, P., Court, R., Mantopoulos, T., Patterson, J., De Santis, M., Cresswell, J., & Mistry, H. (2019). Pembrolizumab for previously treated advanced or metastatic urothelial cancer: An evidence review group perspective of a NICE single technology appraisal. *PharmacoEconomics*, *37*(1), 19–27.

Harty, G., Jarrett, J., & Jofre-Bonet, M. (2018). Consequences of biomarker analysis on the cost-effectiveness of cetuximab in combination with FOLFIRI as a first-line treatment of metastatic colorectal cancer: Personalised medicine at work. *Applied Health Economics and Health Policy*, *16*(4), 515–525.

Hawkins, N., & Scott, D. A. (2011). Reimbursement and value-based pricing: Stratified cost-effectiveness analysis may not be the last word. *Health Economics*, *20*(6), 688–698.

Minton, P., Raiffa, H., & Schlaifer, R. (1962). Applied statistical decision theory. *American Mathematical Monthly*, *69*, 72.

Oliehoek, F. A., & Visser, A. (2010). A decision-theoretic approach to collaboration: Principal description methods and efficient heuristic approximations. In R. Babuška, & F. C. A. Groen (Eds.), *Interactive Collaborative Information Systems* (pp. 87–124): Springer Berlin Heidelberg.

Ondra, T., Jobjörnsson, S., Beckman, R. A., Burman, C.-F., König, F., Stallard, N., & Posch, M. (2016). Optimizing trial designs for targeted therapies. *PLoS One*, *11*(9), e0163726.

Padmanabhan, S. (2014). Chapter 1 - pharmacogenomics and stratified medicine. In S. Padmanabhan (Ed.), *Handbook of Pharmacogenomics and Stratified Medicine* (pp. 3–25): Academic Press.

Sahlin, N.-E., & Hermerén, G. (2012). Personalised, predictive and preventive medicine: A decision-theoretic perspective. *Journal of Risk Research*, *15*(5), 453–457.

Shabaruddin, F. H., Fleeman, N. D., & Payne, K. (2015). Economic evaluations of personalized medicine: Existing challenges and current developments. *Pharmgenomics Pers Med*, *8*, 115–126.

Wilsdon, T. B., Anthony, Edwards, G., & Lawlor, R. (2018). *The benefits of personalised medicine to patients, society and healthcare systems*. Charles River Associates.

Zaric, G. S. (2016). Cost implications of value-based pricing for companion diagnostic tests in precision medicine. *PharmacoEconomics*, *34*(7), 635–644.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.