
Task Force Co-Chairs
Lou Garrison, PhD, Pharmaceutical Outcomes Research & Policy Program, Department of Pharmacy, University of Washington, Seattle, WA, USA
Adrian Towse, MA, Office of Health Economics, London, England, UK

Task Force Members:
Andrew Briggs, MSc, DPhil, University of Glasgow, Glasgow, Scotland
Gerard de Pouvourville, PhD, ESSEC Business School, Cergy Pontoise, France
Jens Grueger, PhD, F. Hoffmann - La Roche AG, Basel, Switzerland
Penny Mohr, MA, Center for Medical Technology Policy, Baltimore, Maryland, USA
J.L. (Hans) Severens PhD, Erasmus University Rotterdam, Rotterdam, Netherlands
Paolo Siviero, BA, AIFA (Italian Medicines Agency), Rome, Italy
Miguel Sleeper, ACMA, Access to Medicines Centre of Excellence, Glaxo SmithKline, Inc., Brentford, UK
Abstract

There is a significant and growing interest among both the payers and producers of medical products for agreements that involve a “pay-for-performance” or “risk-sharing” element. These payment schemes—called “performance-based risk-sharing arrangements” (PBRSAs)—involve a plan by which the performance of the product is tracked in a defined patient population over a specified period of time and the level of reimbursement is based on the health and cost outcomes achieved.

There has always been considerable uncertainty at product launch about ultimate real-world clinical and economic performance of new products, but this appears to have increased in recent years. PBRSAs represent one mechanism for reducing this uncertainty through greater investment in evidence collection while a technology is used within a health care system. The objective of this task force report is to set out the standards that should be applied to “good practices”—both research and operational—in the use of a PBRSA, encompassing questions around the desirability, design, implementation, and evaluation of such an arrangement. This report provides practical recommendations for the development and application of state-of-the-art methods to be used when considering, using, or reviewing PBRSAs.

Key findings include the following observations. Additional evidence collection is costly, and there are numerous barriers to establishing viable and cost-effective PBRSAs: negotiation, monitoring, and evaluation costs can be substantial. Good governance processes are also essential. For good research practice in PBRSAs, it is critical to match the appropriate study and research design to the uncertainties being addressed. The information generated as part of PBRSAs has public good aspects that need to be considered.

The societal desirability of a particular PBRSA is fundamentally a concern as to whether the cost of additional data collection is justified by the benefits of improved resource allocation decisions afforded by the additional evidence generated. The ex post evaluation of a PBRSA will be a multidimensional exercise that considers whether the assessment of projected long-term effectiveness is changed as well as the quality of the evidence generated. PBRSAs should also be evaluated from a long-run societal perspective as investments that affect both utilization and reimbursement arrangements, with implications for dynamic as well as static efficiency.
Introduction

There is a significant and growing interest among the payers and producers of medical products for agreements that involve a “pay-for-performance” or “risk-sharing” element. These payment schemes—called “performance-based risk-sharing arrangements” (PBRSAs)—involve a plan by which the performance of the product is tracked in a defined patient population over a specified period of time and the level or continuation of reimbursement is based on the health and economic outcomes achieved. One database identified 116 cases of these types of arrangements for medicines and other medical products since 1997 [1], with slowly growing numbers in the most recent years. This broad trend represents, in part, a response to the growing cost of new drugs and other innovative medical products and the desire of payers to obtain greater certainty and greater value for the money spent.

There has always been considerable uncertainty at product launch about ultimate real-world clinical and economic performance of new medical products. The uncertainty and concomitant financial risk to the payer for a new treatment that does not work as well in the real world has increased along with the price of the new treatment, whether a biologic, device, or other medical technology. PBRSAs represent one mechanism for reducing this uncertainty through greater investment in evidence collection while a technology is in use within a health care system.

Information about what works in medical care is, in economic jargon, a public good—one person’s use of the information generally does not keep others from using it—whether it is generated by public or private entities. Public authorities who negotiate and fund evidence generating arrangements need to follow good research practices to improve the quality of the information derived and to make the results of that research public where possible. Private insurers, who have less of an obligation for transparency, can still benefit from good research practices as they seek valid scientific answers to the outcomes questions embedded in the arrangements they negotiate. Encouraging them to put their findings in the public domain can generate greater public benefit as well, so long as it does not deter them from agreeing to PBRSAs.

The objective of this task force report is to set out the standards that should be applied to “good practices”—both research and operational—in the use of a PBRSA, encompassing questions around the desirability, design, implementation, and evaluation of such an arrangement. This report provides practical recommendations for the development and application of state-of-the-art methods to be used when considering, using, or reviewing PBRSAs.

Defining PBRSAs

These types of arrangements fall under a variety of names and categories: outcomes-based schemes, risk-sharing agreements, coverage with evidence development (CED), access with evidence development (AED), patient access schemes (PAS), conditional licensing, and managed entry schemes (MES) [2-9]. For
purposes of this discussion, we group all of these under the broad term “performance-based risk-sharing arrangements.”

Under our working definition, a PBRSA exhibits the following key characteristics:

1. **There is a program of data collection** agreed between the manufacturer (or provider, in some instances) and the payer. It can be initiated or required by the payer—to address uncertainties about effectiveness (including possible unintended or adverse consequences), thereby reducing uncertainty about the expected cost-effectiveness of a medicine (or device or diagnostic) in the health care system.

2. **This data collection is typically initiated during the time period following the regulatory approval** (which may be full, adaptive or conditional), but before the full diffusion and uptake of the health technology.

3. **The price, reimbursement, and/or revenue for the product are linked to the outcome of this program of data collection** either explicitly by formula or implicitly through an option to renegotiate coverage, price, and revenue later. The arrangements are primarily concerned with reducing uncertainty about expected health outcomes (usually—but not always about)—better estimates of real-world effectiveness, including any adverse effects). These, in turn, can reduce uncertainty about the appropriate use of the product, and about overall effectiveness and cost-effectiveness, increasing the confidence of payers and health care professionals in using the technology.

4. **The data collection is intended to address uncertainty** concerning one or more of the following:
   - the expected outcomes of the treated population as compared to those obtained with the use of an alternative treatment regime to that in trials or with different combinations or sequencing of new and alternative therapies;
   - the efficacy or effectiveness in a broader, more heterogeneous population than in registration trials or in pre-licensing testing;
   - the effects on longer term or more clinically-significant endpoints than those included in registration trials (which—in the case of a drug—may have used surrogate markers) or in pre-licensing studies (e.g., for procedures or devices):
   - whether health care providers’ management of the patient will change the relative benefits and harms under conditions of usual care;
   - the size and value of cost offsets, such as fewer hospital visits;
   - the proportion of patients who will respond, i.e., achieve a pre-set (minimum) outcome which may be an intermediate/surrogate endpoint;
   - the numbers and types of patients likely in practice to be treated with the new therapy;
• whether the patients treated are the “right” ones, i.e., they have attributes matching those patients which, on the basis of current evidence, the payer is willing to fund.

5. These arrangements provide a different distribution of risk between the payer and the manufacturer than the historical manufacturer-payer relationship. Arrangements that are simply disguised price discounts—and are not concerned with clinical performance—are excluded from this definition. United Kingdom (UK) Patient Access Schemes include a number of these [10]. However, a number of the operational aspects that underpin a successful PBRSA also apply to such schemes.

[See Box 1 for more information on UK PBRSAs.]

From a broader societal perspective, a PBRSA can be thought of as an investment to gather more data to resolve one or more of the above-mentioned uncertainties. The fundamental motivation for a PBRSA is that the manufacturer and the payer hold different views about the potential value of a new intervention or about their willingness to accept the uncertainty around that value. The manufacturer wants a higher price than the payer thinks is justified given the evidence. The payer is concerned about “decision uncertainty”--the probability of paying for a product that is not effective or cost-effective following adoption in their health care system [11].

Investment in a PBRSA should lead to a price that will align the rewards to the manufacturer with the value that the patients—represented by the payers—would assign to the new intervention. To evaluate the investment as such, in theory, we will want to compare the costs of the additional evidence generation with the benefits in terms of making improved resource allocation decisions. The short-term, static efficiency benefits, including ensuring that the new intervention is used in the appropriate population, will be easier to measure than the long-term, dynamic efficiency benefits that are assumed to come from aligning incentives in a way that promotes optimal research and development.

A Taxonomy for PBRSAs

As suggested in the previous section, there are many terms that have been used to describe the types of arrangements considered in this paper. It is helpful to distinguish between those that fit our definition and those that do not. To draw these distinctions, we are able to build upon previous taxonomies that have been published in the literature [6,7,12,13]. For example, McCabe et al. (2010) recognize that such schemes do not always include a research component, making an important distinction between population- versus patient-level schemes, the latter characterizing a situation where value for money is guaranteed without the need for subsequent review of the reimbursement decision [6].

Another taxonomy developed by Carlson, et al. (2010) was based on an inventory of published schemes categorized in terms of timing, execution, and health outcomes [7]. This taxonomy made a clear distinction between schemes that are health outcomes-based and those that are not. Towse and Garrison (2010) in their taxonomy also made the distinction between outcome-based and non-outcome-based, and
also between (i) those with agreements that specified how evidence would be translated into revisions to price, revenues, and/or use and (ii) those that instead specify an evidence review point where renegotiation would occur [12]. They also make the point that the outcome evidence either can come from a randomized trial on a subset of patients (who may not necessarily be in the same health system) or through an observational study of the patients being treated. They also draw a distinction between (i) uncertainty about the performance of the drug within a subgroup of patients and (ii) uncertainty as to the subgroups of patients who would in practice receive the drug.

For purposes of defining good research practice, we distinguish between those payer-producer/provider arrangements that measure health outcomes in characterizing performance and those that focus on sharing costs (non-health outcomes). We consider the latter “cost-sharing arrangements” only, which fall outside our definition of a PBRSA. Examples of such arrangements are budget- or utilization-capping, variable or fixed discounts, and price-volume arrangements not linked to the underlying cost-effectiveness of treatments in different patient subgroups [14].

For PBRSAs that measure health outcomes, we follow McCabe et al. (2010) and further distinguish between those that attempt to directly manage utilization and guarantee cost-effectiveness (i.e., “utilization-based”) versus those that include a strong research element (i.e., “research-based) [6]. The former, performance-linked real-world arrangements, have the primary objective of assessing utilization (for which patients are treated) and/or patient level outcomes (has a target outcome been achieved for a patient?). Such schemes adjust payments and prices in an attempt to ensure value for money. On the other hand, the research-oriented arrangements are focused on covering the procedure for a period of time in order to develop further evidence that will reduce the uncertainty about the long-term outcomes expected to be achieved in groups of patients instead of individual patients.

Drawing upon the previous taxonomies, Figure 1 depicts these distinctions. From this figure, it can be seen that PBRSAs are a specific group of schemes among all possible payer-producer/provider arrangements. Pure “cost-sharing arrangements” shown in gray on the left-hand side are not within our definition.

The taxonomy separates two types of PBRSA schemes: those whose goal is provide coverage while the evidence is developed and those whose goal is to manage utilization and control the cost-effectiveness of a new technology in the real world. CED is a payer-producer (or producer-provider) arrangement that attempts to reduce decision uncertainty regarding coverage policy. In short, such schemes link population-level payment or reimbursement to prospective data collection. They can differ according to the number of patients “exposed” to the technology.

For example, all new patients might be treated using the new technology (“only with research” or OWR) or only those patients (“only in research” or OIR) included in a registry or trial [15]. We include both within the term CED as forms of PBRSAs. See Walker et al. (2012) for a discussion of the criteria that might lead to the use of either an OWR or OIR scheme [16]. CED schemes can exist either with a pre-specified agreement or without. PBRSAs without a pre-specified agreement as to how the results of the
research will be used to adjust price, revenues, or coverage will require renegotiation at a later time [17,18, 19, 20]. [See Box 2 for more information on CED in the US.]

In contrast to these CED arrangements, the other major category of PBRSAs is of those that aim to manage utilization and guarantee the cost-effectiveness of a new technology in the real world. In principle, such schemes link performance at the individual patient level to payment or reimbursement for a new technology. In some schemes, payment is related to process of care. This means that reimbursement is specified \textit{ex ante} to depend on the clinical decision-making process, e.g., a provider’s compliance with clinical guidelines or selection of individual patients based on a biomarker, such as a genetic test.

Figure 1. PBRSA Taxonomy

Other schemes focus on \textit{ex post} reimbursement, measuring intermediate or clinical endpoints. These arrangements include (i) "outcomes guarantees"—meaning payment for responders only—or (ii) "conditional treatment continuation"—meaning payment for continued use of the product based on intermediate endpoints. Thus, in contrast to CED, reimbursement in these schemes is generally determined for each individual patient, either retrospectively or prospectively by rule. Nevertheless, sometimes the data
collected in such real world-utilization schemes could be aggregated to the population level to adjust overall payments. In addition, for research purposes, post hoc analyses to support additional population-level decision-making are a possibility if the patient-level data have been appropriately collected.

**GRPs for PBRSAs: Overview of Key Good Practice Questions**

For those PBRSAs whose goal is to provide coverage while the evidence is developed, either the payer or both the payer and the manufacturer (or service provider) together will have to address four research-related, good practice questions concerning: (Q1) the desirability of the PBRSA (as opposed to some other form of reimbursement or research arrangement), (Q2) research design, (Q3) implementation approach, and (Q4) evaluation method. In some instances, the payer and manufacturer (or service provider) may have to reach a formal legal agreement as to why the PBRSA is desirable and under what conditions it will move forward. In others, the payer may unilaterally decide to delay approval and collect additional data, or recommend or require another party to collect it. For those PBRSAs that aim to manage real world utilization and guarantee the cost-effectiveness of a new technology, the four good practice questions are also relevant, although only in part.

**Q1. Desirability: Is a PBRSA the appropriate way forward given the uncertainty and the alternative methods to reduce this uncertainty?**

*When is a PBRSA worthwhile?*

The key issue is that at product launch the payer could have considerable uncertainty as to whether the product or service offers good value for money. In theory, a payer has four major options [21]:

1. Adopt (or partially adopt) despite the uncertainties and wait for more information to revisit the decision.
2. Refuse to adopt until the manufacturer supplies better evidence to address the uncertainty.
3. Demand or mandate a lower price such that the uncertainty about value is reduced.
4. Enter into a PBRSA that (a) manages utilisation/outcome at the patient level or (b) is a form of CED where evidence is collected across patients for a review, potentially leading to pre-specified adjustments.

Each of these options is associated with costs and benefits. A value-of-information (VoI) framework—comparing the costs of additional data collection with the benefits of improved resource allocation decisions—should be used for weighing these options [22]. The general question is whether a PBRSA can effectively and efficiently address the uncertainties that remain following marketing authorization.

Issues that would need to be considered include:

- The expected value(s) of research options in terms of reduced uncertainty for the decision maker.
- The value of any additional research (and, therefore, of a CED or research-based PBRSA) will depend in part on the price of the product and its expected use in the patient population;
- the cost (in terms of out-of-pocket costs and administrative burden) of collecting evidence;
• the cost (in terms of expected health loss) of any delays in access that might result from use of a scheme;
• the existence of any irreversibility in the process—for example, if an adoption decision would make subsequent withdrawal of the product much more difficult or make research less feasible or even impossible in some circumstances [16]. Thus, a PBRSA may enable research to take place alongside use (OWR) or, in other cases, the PBRSA will restrict this to OIR.

A utilisation-based PBRSA will change expected cost-effectiveness by changing the effective price or use of the product through, for example, an outcome guarantee or the use of an intermediate response as part of conditional treatment continuation. There will be implementation costs associated with such a scheme. The relative merits of research-based PBRSAs and utilisation-based PBRSAs will need to be assessed alongside the other options of either (a) adoption without any expectation of further evidence collection, (b) a refusal to adopt until further evidence is available, or (c) adoption at a lower price.

**Desirable for whom?**

When a manufacturer and a payer are in negotiation on coverage and reimbursement, a manufacturer considering a PBRSA will have to weigh the pros and cons of the additional complexity and cost of the PBRSA against alternatives, for example, offering an upfront price reduction [13]. But to project complexity and cost, the manufacturer will also need to address questions Q2-Q4: evidence collection, implementation, and evaluation.

In these instances, whether to propose or accept a PBRSA will be a business decision for both the manufacturer and the payer. When proposing a PBRSA, the manufacturer should have established that a valid and efficient process of evidence collection following good professional practice is feasible. The payer should be realistic regarding the level of uncertainty as well as the cost of data collection and implementation of the scheme when requesting a PBRSA. Neither party is per se subject to a discussion of good research practices, although a public payer may indicate scientific research criteria that have to be met. In any case, the quality of research needs to be high enough to ensure the resulting data are robust enough to address the key uncertainties.

PBRSAs can be desirable from the manufacturer’s perspective because they may be a way to overcome payers’ aversion for risk and reduce time to market access. If the alternative is extending clinical development time prior to launch, this risks (i) still not demonstrating a significant improvement in effect, (ii) reducing the period of exclusivity associated with patent protection, and (iii) enabling a competitor to come into the market with similar or better results before the manufacturer has entered the market.

For a payer, denying reimbursement on the basis of high uncertainty opens them to the risk of depriving patients of the benefits of a new medical product. On one hand, they may thus face political criticism for any loss of societal welfare if the product is later found to be cost-effective. On the other hand, they face criticism for any losses if the product is not cost-effective and/or provides little or no benefit to patients.
It is worth noting that these perspectives may not be symmetric: payers may lose less in the short term by forgoing an agreement. High uncertainty with regard to effectiveness may easily justify denial of reimbursement or severe restrictions on indications, which is politically sustainable and may be seen as financially prudent. PBRSAs may well seem more desirable for industry than for payers. Yet, the use of a PB RSA (as opposed to a denial) may be socially optimal, as well as in the commercial interest of the manufacturer. [See Box 3 on the post-launch studies in France.]

PBRSAs may also lead to the pursuit of opportunistic strategies by either manufacturers or payers. On the manufacturers’ side, they might define their price target according to the anticipated risk of not meeting the target. On the payers’ side, they might, if they have bargaining power, set their reservation price on the basis of the probability of a failure of the product to meet its target.

Finally, it is important to remember that the payer is the agent of patients: good PBRSAs are only good if they make patients as a whole better off in terms of health outcomes.

Q2. Evidence collection: Which PB RSA research design is most appropriate to collect evidence that addresses the relevant uncertainties?

The answer will depend on the uncertainty that the PB RSA evidence collection is trying to address:

- Uncertainty around whether the medical product or service will be used in the right patients, which may be because not all patients will respond, or because effectiveness and cost-effectiveness differs across indications or patient groups.
- Uncertainty at launch around clinical or economic outcomes (effectiveness vs. efficacy, final outcomes vs. surrogates, or about size of cost offsets).

We enumerated the many possible sources of uncertainty above. However, the preferred study design may differ for questions such as the optimal subset of patients in an indication versus questions about the transferability of an efficacy result to effectiveness in the real world.

GRPs for evidence collection in PBRSAs should build upon previous GRPs for specific types of studies. Previous ISPOR Task Forces have defined GRPs in a number of relevant areas: modelling, nonrandomized studies of treatment effects, RCTs, registries, and prospective observational studies [23-27]. Although ISPOR and others have developed GRPs around a wide range of study designs, much less work has been done linking particular designs to specific research questions. This linkage or translation is an active area of research, particularly in the field of comparative effectiveness research [28].

The number of general options for research design to collect data post-launch is fairly limited and reasonably well defined. These include:

- a traditional targeted randomized clinical trial (RCT), focusing on efficacy;
- a large pragmatic clinical trial (PCT), randomized but with less rigorous entry inclusion or exclusion criteria;
- a prospective observational study of patients without randomization;
- a hybrid design that includes observational cohorts and retrospective data;
Each of these designs has strengths and weaknesses in the context of PBRSAs. Table 1 summarizes the factors that favour a randomized versus observational design. What is important is that the study is designed to answer the question at hand. Historically, there is a general consensus that properly sized RCTs are the strongest method for determining a treatment effect. However, they can be criticized for limited generalizability.

Large PCTs can be costly and difficult to manage, but can offer a better estimate of real-world effectiveness. Prospective observational data collection, for example, using registries, can be useful for estimating real-world effectiveness as well as the relationship between surrogate endpoints and long-term outcomes. [See Box 4 for more information on PBRSAs in Italy.]

These studies can also be comparative by using historical or matched controls. Methods such as propensity scoring can be used to correct for the likely selection bias. Retrospective study designs can be helpful for measuring historical controls or relationships between surrogate endpoints and long-term outcomes. They are also subject to selection bias.

Table 1. Factors Affecting Selection of Randomized vs. Observational Design

<table>
<thead>
<tr>
<th>Factors Favoring Randomized Design:</th>
<th>Factors Favoring Observational Design:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Feasible to randomize</td>
<td>• When relative effectiveness is of interest: Effect size is relatively large and/or selection bias can be reasonably controlled</td>
</tr>
<tr>
<td>• Risk of getting the answer wrong is large (large impact on mortality, resource use, or treatment patterns)</td>
<td>• Relatively rare adverse events, safety is most important outcome of interest</td>
</tr>
<tr>
<td>• Prognostic variables are unclear; most variation in outcomes is unexplained</td>
<td>• Major focus is on adherence or compliance with therapy</td>
</tr>
<tr>
<td>• Biologic process of disease is not well understood</td>
<td>• Interest in associations of outcomes by patient subgroups, observed practice patterns</td>
</tr>
<tr>
<td>• Modest anticipated differences in effect size</td>
<td>• Risk of getting it wrong is low</td>
</tr>
<tr>
<td></td>
<td>• Cost or resource use is main interest</td>
</tr>
</tbody>
</table>

Adapted from Gliklich (2012) [28].

Under randomized design, there is a continuum from more explanatory to more pragmatic designs. For assessing comparative effectiveness and cost-effectiveness, the more pragmatic designs are preferred, including: the parsimonious use of inclusion/exclusion criteria; clinically meaningful outcomes; protocols that interfere minimally with practice patterns; and usual care settings.
Observational studies, such as registries or post-marketing surveillance cohort studies, could be feasible to monitor effectiveness and safety of investigated medicines in clinical practice. After the period of the study observation, results could be modeled in order to investigate if the continued use of a PBRSA is cost-effective. An observational study linked to a monitoring registry might be a feasible approach to undertake.

Although most PBRSAs to date (with the exception of those in the US) have not been based on an RCT design, it is an important option. An RCT can be designed to clarify an aspect of efficacy or explore effects in a key subgroup of patients or in the validation of a biomarker. Arguably, this is the basis for the Temporary Authorization for Use—ATU—scheme in France. If it is not possible to conduct an RCT alongside use of the product by the payer, then the RCT could either be conducted in an OIR scheme or in an OWR scheme in which the PBRSA relied on RCT data collected in another jurisdiction, providing any necessary adjustments for the transferability of data are made [29].

As Hutton et al. (2007) and others have emphasized outcome measures should be selected with care [3]. They should be clear, measurable, objective, realistically achievable and relevant. All parties will need to ask:

- Is the type of evidence being promised sufficient to address the uncertainty in quantitative/qualitative terms?
- Are the endpoints the desired outcomes (or acceptable surrogates)?
- In the case of CED schemes, is the duration of study sufficient to deliver a result against the measures, and is the study population sufficiently representative?

Choice of the outcomes for PBSRAs will also be influenced by the scope of regulatory approval, especially for drugs. If the manufacturer wishes to use the PBRSAs results for promotion, any claims need to be consistent with the label.

For efficiency purposes and in order to answer questions around effectiveness, the design will often make use of extensive data collected routinely (e.g., claims data, lab and pharmacy data, hospital and electronic medical records). But prospective study designs for PBRSAs that rely on these records present a number of issues and challenges in terms of implementation, which are discussed in the following section.

**Q3. Implementation: How should a PBRSA be implemented, governed, and reported?**

**Implementation**

Aspects of good implementation follow from clarity on the desirability of using a PBRSA and on the type of evidence being collected as part of the PBRSA. These aspects have been addressed in the literature, notably in the 2009 Banff Summit principles (Menon et al., 2010) and by Hutton et al. (2007) [30,3]. They include the following:

- *Is the scheme measuring appropriate outcomes?* PBRSAs should be clinically robust, clinically plausible, appropriate, and monitorable. For example, if the scheme is based on a patient response,
there must be a relatively straightforward way to measure a patient’s clinical response. Standard procedures for reporting and analysis of adverse events will need to be followed.

- **Are the costs acceptable?** The cost to the commissioning body or healthcare system arising from the PBRSA should be proportionate to the potential gains. We can note requirements expressed in the UK PPRS (2009) which related to both CED and utilisation schemes [10]:
  - When considering the burden, the full cost to the health care system of any patient access scheme should be included in the costs considered by the commissioning body.
  - Any scheme should be operationally manageable and without unduly complex monitoring, disproportionate additional costs and bureaucracy.
  - Any burden for the health care system should be proportionate to the benefits of the scheme for the NHS and patients.

- **Is the time horizon realistic?** A PBRSA process should set clear target dates by which the future contingent access decision will be made. Although this can vary by disease, Hutton and colleagues suggest that a PBRSA study period longer than three years faces the risk of becoming increasingly irrelevant in the face of changing clinical practice and technological advancement [3]. Even so, difficulties can easily arise if, for example, patient recruitment to studies is slower than planned. They further explain that if a PBRSA continues indefinitely, without the benefit of new evidence, it merely replicates the unsatisfactory situation that gave rise to it in the first place (i.e., coverage with inadequate evidence). It is important then that all participants of the PBRSA make sure that collection of relevant new data can be accomplished within a realistic period before entering into a scheme.

- **Are the funding arrangements clear?** There is a general presumption that manufacturers and sponsors of technologies will finance extra data collection but is not always be the case. There are examples where a government has agreed to fund arrangements (e.g., Medical Services Advisory Committee (MSAC) in Australia, Catalan Agency for Health Technology Assessment (CAHTA) in Spain). The United Kingdom Multiple Sclerosis Risk Sharing Scheme (UK MS RSS) was jointly funded by the DH and the companies involved [31,32].

- **How is responsibility for undertaking data collection and analysis allocated?** In a CED scheme it should be subject to normal research governance arrangements as we consider in the next section.

- **What will be the process for analysis and review of the evidence to make a revised decision on price, revenue or coverage?** A PBRSA should have a process in place to underpin a “Decision with Further Evidence” [3]. In some agreements there will be a pre-agreed link between the evidence and the revised decision (e.g., an adjustment to price to ensure a cost-per-QALY of a certain amount) subject to any arbitration and appeal arrangements. In others, use of the analysis to change price, revenues, or coverage will be subject to negotiation.
• Will discounts or rebates be paid during the course of the scheme (for example based on provisional results)? We can note than in utilisation schemes such as responder-based reimbursement this is integral to the scheme.

**Governance**

Some PBRSAs are bilateral commercial agreements between a private payer and a manufacturer, such as the case between Merck and Cigna for selected anti-diabetic drugs in the United States [33]. In these situations, a formal governance structure is not essential. Likewise, in the case of arrangements between a public payer and a manufacturer for utilisation schemes (as in Italy [Box 4] or the UK [Box 1]) or where an agreement between the public payer and the manufacturer involves a price adjustment linked to the outcome of research within an existing formal structure for such arrangements (as in the case of the UK CED OWR Patient Access Scheme for Votrient® (pazopanib).

In other cases, PBRSAs involve agreements among multiple stakeholders. For example, in the United States, the creation of the implantable cardiac defibrillator (ICD) registry in support of Medicare’s 2005 CED decision involved a partnership among professional associations, public and private insurers, and federal sponsors of clinical research, hospitals, a quality improvement organization, and others [17]. The ICD registry is managed by the American College of Cardiology. Funding for the registry is sustained through fees levied on participating hospitals. Research funding came from a variety of sources including the American Health Insurance Plans, the National Institutes of Health, and the Agency for Healthcare Research and Quality. Medicare funding pays only for the implanted devices.

The UK MS RSS was led by a Steering Group that included the four sponsors of medicines concerned, the Association of British Neurologists, the Multiple Sclerosis Society and Trust, the Royal College of Nursing, and the Association of Multiple Sclerosis Nurses [34]. Those PBRSAs that have multiple stakeholders and/or the involvement of taxpayer funding of research as well as payment for the intervention should be governed through a diverse steering committee that includes patients, manufacturers, disease advocacy groups, professional associations, and other major stakeholders, with an outlet for receiving broad public comment.

In these circumstances it is essential that there is a formal governance structure in place to ensure transparency of the nature and the aims of the scheme, accountability and mitigate conflicts. Transparency about any CED pricing arrangement itself is not essential and may be deleterious: transparency about process is much more important. Many schemes have failed to date when the original aim, research design, or data collection was changed as a result of competing aims in the political arena among stakeholders [35,32]

Effective governance will require the following:

- The governance committee should have a clear charter specifying who is involved and their respective roles.
• The governance agreements should specify the aims of the PBRSA, who has access to data, who can publish, the process for vetting manuscripts, and the final steps for managing and disseminating the research—that is, the stopping rules for data collection, and how the results will be used.

• Funding arrangements should be clearly specified up front with clear information about the types of information or products each funder will receive as a result of their involvement.

• The agreements should spell out a process to ensure data quality, including the conduct of regular audits.

• All conflicts of interest should be declared and to avert undue political influence over the outcome, it is desirable to have a process in place for the independent review of research designs and the neutral, independent conduct of research to achieve the clearly stated initial aims of the scheme.

**Reporting of Results**

Most PBRSAs to date involve a public payer. One could argue that PBRSA contracts between private payers and manufacturers have a different status in terms of public reporting of results. Recently, for both ethical and legal reasons, both private and public entities that sponsor prospective RCTs, have assumed greater public reporting requirements, e.g., they have to report whether an intervention “worked or not.” However, from an evaluation standpoint, because the reporting obligations or practices vary between public and private payers, the types of information available to third parties for any evaluation are likely to differ.

Given that the evidence generated would be a public good, there would seem to be some implicit (and sometimes explicit) requirements for the disclosure of such public information. When a private payer and manufacturer reach an agreement, there may not be clear reporting requirements in terms of what must be made public. In fact, historically, most of these transactions were viewed as confidential and proprietary.

However, in recent years, with the advent of public reporting requirements, such as clinical trials.gov, there seems to be a growing recognition that these activities involve the voluntary participation of patients and there is some ethical obligation to report on research results—at least for prospective clinical trials. Depending on the extent of this reporting requirement, the reporting differences between PBRSAs with public payers and private payers are lessening, although they are still substantial.

The public goods aspects of the information generated by PBRSAs should not be overlooked. The incentives for both manufacturers and private payers to “free ride” is always there, and some public payers may want to free ride on other public payers (e.g., in other countries) [2,36-8]. From a global (societal) perspective, new PBRSAs that follow GRPs may need to be subsidized or encouraged in order overcome these disincentives.

**Q4. Evaluation: Has the PBRSA achieved its objectives? Was it good value from a health system perspective?**
This question links back to expectations /assumptions in Q1, which can be addressed from multiple perspectives—manufacturer, payer, patient, provider, and society. We can ask several different questions:

Are we more knowledgeable about the technology in question? Have patients benefited from access? How do the costs of the scheme relate to the value of the benefits? And so on.

A comprehensive evaluation will therefore need to consider multiple perspectives. Certainly, patient, provider, manufacturer, and payer satisfaction with the scheme will need to be assessed. From a societal perspective, a PBRSA can be viewed as an investment decision to gather more data about product performance, and can be appraised against alternatives. Where a health system seeks to maximise the health of its population from a fixed budget (as for some private insurance schemes or national health schemes, like the UK), then any appraisal of a PBRSA has to take into account the opportunity cost from a health care provider/funder perspective.

Some of the additional costs of the data collection, including negotiation, monitoring, and assessment, can be measured relatively easily. It will be more difficult to measure costs, such as the foregone benefits to patients not receiving access during the data collection if the design is based on an OIR scheme. For example, a CED scheme with later renegotiation could mean both use restricted to a subpopulation and a price change. The result could be more efficient by matching societal willingness to pay to the appropriate subgroups, sending manufacturers a clear signal about societal willingness to pay, as well as restricting use to the appropriate population.

A full assessment will also need to consider reversal effects, i.e. whether the evidence collected led to changes in the use of the technology or in its price. It may be difficult to change provider and patient behaviour in response to a subsequent restriction in use, and a change in price may be resisted by manufacturers (if it is downwards) or payers (if it is upwards).

Although a PBRSA scheme (either a utilisation scheme or a CED scheme) may have been developed with the principles of value of information in mind (that the benefits associated with generating additional evidence exceed the costs of generating that evidence) there are challenges in using these techniques at the evaluation phase in the case of a CED scheme. Ultimately, it is impossible to assess the ex post value of information generated for a single CED scheme. [See Box 5 for a discussion of this issue].
Were the intended outcome measures collected?

Was uncertainty in associated parameter estimation reduced for the outcomes that were the focus of the scheme?

Did the scheme run to budget?

Was the integrity of the design/estimation maintained?

Did the governance arrangements work?

Did the process to underpin a decision with further evidence work?

The last point is very important. Ultimately, of course, in order to meet the objectives of the PBRSA scheme, it is necessary to show that the decision-making following the reporting of the scheme is improved. This success requires more than just the ability to show the process indicators above. Where an utilisation scheme is being run, appropriate decision-making will require the ability to show that the agreed outcome adjustments were made to guarantee the cost-effectiveness of the intervention.

Where an OWR scheme was implemented with a further review, then it will be necessary to demonstrate that the appropriate decision was made in light of the evidence generated: this could be a “no change” decision where appropriate, or a reversal of previous recommendations, or in the case of a price negotiation, that the price was adjusted. Where an OIR scheme was implemented; did the generation of additional evidence lead to a confident recommendation (either positive or negative) for the whole patient group following the completion of the scheme?

Pueg-Peiro et al. (2011) conducted a systematic literature review to identify existing knowledge about the costs and benefits, assessed either quantitatively or qualitatively, of PBRSAs [38]. They found that more than 40 per cent of the publications referred to the UK MS RSS, and no studies were able to evaluate the overall economic impact of a scheme. All studies included qualitative discussions of costs and benefits, with the exception of the UK MS RSS where some costs were reported. Neumann et al. (2011) reviewed five PBRSAs in the US and UK and conclude that they are hard to implement in practice [40]. The results from Italy and other EU countries are also unclear and the schemes are in evolution [41-3]. [See Box 6 for more information on the PBRSA experience in the Netherlands.]

The experience of CED in the US is covered in Box 2. The main form is CED using OIR or OWR with no explicit agreement between the manufacturer and the payer, but an implicit assumption that the data will be used for future coverage decisions. Difficulties have emerged around study financing and design, but most importantly around decision making with further evidence.

In the case of the UK MSS, the results of scheme are discussed in Box 1. Issues were raised as to: the design of the study and the time delays in generating the evidence; the enforceability of the contract in relation to the link between prices and outcomes; problems of governance of the scheme [44]. Such evidence as we have on the other non-US CED OWR scheme, that for bosentan (Tracleer®) scheme in Australia, suggests it was successful. The registry was populated. The results suggested some reduction in
price was appropriate. A competing product was listed at a 15% price discount and Actelion agreed to a PBAC request for the same discount [45].

Although the UK PASs include discount arrangements as well as PBRSA utilisation and OWR schemes, the experience is relevant. Williamson (2010) reports on a survey of oncology pharmacists in NHS hospitals [46]. Transaction costs for the NHS were the biggest concern. However, a review by the DH focussed on the additional numbers of patients receiving access to drugs deemed cost-effective by NICE (including transaction costs) [47]. The Italian PBRSA utilisation schemes appear to have been well received. This may reflect in part use of a national electronic patient registration system reducing transaction costs (see Box 4).

Overall, however, the literature suggests there is an important gap in structured ex post evaluation of PBRSAs. Utilisation schemes appear to have been more successful to date than CED schemes. However, the evidence is limited, mixed, qualitative, and partial.

Conclusion

This task force report has reviewed the issues associated with defining good research and operational practices for PBRSAs. Previous analysts, commentators, and task forces have identified, discussed, and addressed many of these issues. Previous ISPOR methods task forces and other professional organizations have defined good research practices for the main relevant study designs. Our intention is to move the discussion forward by defining the scope of the problem and identifying the issues with greater clarity.

The major messages of this report are:

- PBRSAs are an understandable and logical response to increasing pressure for greater evidence of real-world effectiveness and long-term cost-effectiveness for new medicines and other health technologies in the early stage of adoption and diffusion.
- In some cases, PBRSAs manage utilization and reward for performance at the patient level, tracking what happens to each patient. In contrast, the other major type of PBRSA is research-based and uses CED with research, either alongside product use (OWR) or product use only in the context of research (OIR).
- PBRSAs using CED can use observational studies or RCTs in either an OWR or an OIR setting. Where an RCT is preferred but may not be feasible with OWR, then evidence can be collected from an RCT in another jurisdiction, providing the results can be translated into the setting of interest to stakeholders.
- All PBRSAs, including those tracking patients, can provide valuable evidence that is a global public good. The value of that evidence will be enhanced if good scientific practices are followed in research design, implementation, and evaluation.
• Additional evidence collection is costly. It is critical that it is designed to address the main uncertainties that are making payers reluctant to reimburse or recommend use of the product. GRP requires matching the appropriate study design to these uncertainties.

• There are numerous barriers to establishing viable and cost-effective PBRSAs: negotiation, monitoring, and evaluation costs can be substantial. A process to underpin decision-making with further evidence is key. Good governance processes are also essential.

• Because they can generate evidence which is a public good, PBRSAs are likely to be underutilized. This tendency can be countered if public payers assume some greater responsibility for the greater (global) societal welfare and private payers and manufacturers are constrained or incentivized to utilize these agreements when appropriate.

• The societal desirability of specific PBRSA is fundamentally a VoI question, comparing the societal costs of additional data collection with the societal benefits of improved resource allocation decisions. However, the evaluation of the success of a PBRSA should be a multidimensional exercise that tracks not only whether uncertainty about the projected long-term effectiveness is reduced, but also the quality of the research process and evidence generated. It must look at the impact on decision uncertainty and whether there was an effective process to support decision-making with additional evidence.

• PBRSAs should be evaluated from a societal perspective as investments in collecting further evidence to inform both utilization and pricing arrangements, with implications for dynamic efficiency (eliciting the optimal amount of innovation as well as static efficiency (using current resources wisely).

• There is an important gap in the literature of structured ex post evaluation of PBRSAs. Utilisation schemes appear to have been more successful than CED schemes. However, the evidence is limited, mixed, qualitative, and partial. The ability to run successful PBRSA schemes using CED will provide an important additional tool for increasing efficiency in health care. Robust evaluation of such schemes will be important for learning.
Box 1: United Kingdom

Which entities are involved in the process?
The UK has “flexible pricing” and “Patient Access Schemes” (PAS). Both are defined by the Pharmaceutical Price Regulation Scheme (PPRS) of 2009. Under flexible pricing (FP) companies can apply to increase their price if the evidence supports it. It is agreed that NICE will use its normal evidence standards and the cost-effectiveness threshold it used earlier when initially agreeing to use the drug [10]. However, no more detailed arrangements are outlined, and there are no requirements for resubmission. By contrast, PASs are agreement-specific. However, most are “financial” arrangements intended to provide the UK NHS with effective discounts from list price rather than being linked to “outcomes.” The UK PAS, therefore, include pay-for-performance agreements but are mainly types of discount agreements.

What is the general approach and experience in the UK?
Examples of UK schemes include:
- The dose-capping agreement that NICE entered into over ranibizumab (Lucentis®) for macular degeneration could be seen as an effective price discount. Cost-effectiveness to NICE was only acceptable if the NHS paid for up to 14 injections per eye of eligible patients. Novartis will bear the costs of treatment beyond this [48].
- NICE recommended ustekinumab (Stelara®) for severe plaque psoriasis on condition that Janssen-Cilag ensures the costs of treating patients weighing more than 100 kilograms will be no more than patients weighing less than 100 kilograms [49]. This equates roughly to purchasing two vials of ustekinumab (Stelara®) for the price of one.
- The bortezomib (Velcade®) ensures identification of responders. There is retrospective payer reimbursement for non-responders. Responders receive further doses of the product.
- Votrient® (pazopanib) involves a discount on Votrient’s list price to achieve price parity with Sutent® (sunitinib). The NHS will get a financial rebate if Votrient proves inferior to Sutent with regards to its efficacy, in ongoing head-to-head trials. The results of the study will not be known until mid-2012 and the exact scale of the potential rebate has not been disclosed.

Evidence on the costs and effects of schemes is limited. Although the UK PASs are largely discount arrangements rather than pay for performance, the experience is relevant. Williamson (2010) reports on a survey of oncology pharmacists in 31 NHS hospitals. Transaction costs for the NHS were the biggest concern [46]. Variation between the administrative requirements of different schemes added to the problem. There was a concern that, in some cases, money due back may not have been claimed. In other cases, the money came back to the provider hospital but the purchaser (commissioner) was not aware of this. The “two schemes linked to a measurement of clinical response, cetuximab [Erbitux®] and bortezomib [Velcade®], showed a trend towards being the worst. Response-based schemes pose challenges for tracking patients and ensuring claims are made to refund non-responders.” [46; p111].

An example of a PBRSA in the UK
The UK multiple sclerosis (MS) drugs risk-sharing scheme (RSS) addresses outcome uncertainty with an observational study of patient health status with price linked to a cost-per-QALY threshold. The UK MS RSS was negotiated in 2002 between the UK Department of Health and four pharmaceutical companies supplying MS drugs following NICE’s rejection of any use of these drugs by the NHS. It is a 10-year observational study with a historic cohort as a control. It took three years rather than the expected 18 months to recruit 5000 patients at 73 centres. The results of the two year assessment of accumulated disability of the 5000 patients recruited found were not reported until 2009, seven years after the agreement to have a scheme.

In reporting the results, Boggild et al. (2009) said that “the outcomes so far obtained in the pre-specified primary analysis suggest a lack of delay in disease progression” [31]. However, prices were not adjusted downwards on the grounds that the evidence was not conclusive. This raised issues as to: the design of the study and the time delays in generating the evidence; the enforceability of the contract in relation to the link between prices and outcomes; problems of governance of the scheme including the independence of the Scientific Advisory Group (which was vigorously defended by its chair [50]; the usefulness of the Expanded Disability Status Scale (EDSS) as the outcome measure; and the impact on the choice of comparator when evaluating subsequent new drugs for the same indications.
Box 2: United States

What is the general approach and experience in the US?

The process for deciding when to do a PBRSA in the US by both public and private payers has been opportunistic and ad hoc to date. Within the Medicare program, this is largely due to two important barriers to developing a cohesive approach to coverage with evidence development (CED): unclear statutory authority and the lack of a dedicated funding source.

To date there have been more than 20 documented PBRSA initiatives in the US [17]. These initiatives largely focus on devices (8) and surgical procedures (6). Only four of these initiatives focused on drugs. Five were initiated by private industry and the remainder by public or private payers. The main form is CED with no explicit agreement between the manufacturer and the payer, but an implicit assumption that the data will be used for future coverage decisions.

Unlike the experience in much of Europe, over half of US PBRSAs have been randomized controlled trials where access is granted OIR. Data collection has been supported through diverse public and private sources with the majority of clinical trials funded by the main federal clinical research body—the National Institutes of Health (NIH). Complex public-private partnerships have developed to govern and support the costs of establishing and maintaining prospective registries used for OWR schemes.

Which entities are involved in the process?

Private insurers were early innovators with the concept of performance-based risk sharing, exploring a variant then known as conditional coverage in the mid-1990s (fitting into our taxonomy as OIR). Blue Cross Blue Shield plans with enrollees in the Federal Employee Health Benefits Program (FEHBP) began a collaborative demonstration project examining the use of autologous bone marrow transplant for treatment of three diseases: metastatic breast cancer, epithelia carcinoma, and multiple myeloma [51]. NIH-funded trials demonstrated the risks outweighed the benefits for treatment of metastatic breast cancer and leading to a rescinding of coverage for that indication. However, FEHBP members successfully sued to gain access to the procedure outside of the clinical trials. This precedent has made private payers cautious about mandating participation in a randomized clinical trial in order to obtain coverage.

There have been few subsequent attempts at PBRSAs in the private sector in the US. None of these have taken the form of OIR. There have been four documented performance-based outcomes or process guarantees between drug or test manufacturers and private payers since the late 1990s. These arrangements examined lipid-lowering drugs, diabetes drugs, osteoporosis drugs, and a gene expression profiling test used to identify potential responders to chemotherapy for breast cancer [33]. Conditional coverage (only with research) was also offered by one payer in an agreement with a device manufacturer to examine the long-term durability of interventional procedures for treatment of uterine fibroids [17].

The group most commonly associated with PBRSAs in the United States is the Medicare program, which initiated a “Coverage with Evidence Development” program [52]. Medicare distinguishes between policies where coverage is provided only to study participants (coverage under study protocol) and those that offer broad access but require additional data collection (coverage with appropriateness determination). This distinction is manifested largely as policies that require further randomized studies and those that mandate participation in a registry. Since 1995, the agency has issued 13 CED policies, all but two have been coverage under study protocol and most have targeted devices or diagnostics.

Examples of PBRSAs

Notably, Medicare has used the evidence generated from CED policies to inform subsequent coverage determinations in only three occasions. These three cases offer insights into the structure of Medicare CED initiatives and some of the financial, legal, and operational barriers the agency has faced in their successful implementation. The first Medicare CED initiative was the provision of temporary reimbursement for lung volume reduction surgery for emphysema treatment only for beneficiaries who participated in a clinical trial. This initiative had a dramatic impact on treatment patterns [53]. This well-designed NIH-funded trial found that apart from a small subgroup, the surgery potentially increased the risk for mortality and offered little improvement in quality of life. Even with this evidence, Medicare extended coverage for this procedure to all beneficiaries. Yet, the number of procedures dropped dramatically, as physicians responded to the trial evidence. This example demonstrates the difficulty of rescinding coverage once it is offered provisionally through CED.

Uncertainty about the diagnostic benefit of positron emission tomography (PET) for cancer diagnosis, staging, and monitoring despite growing pressure from the clinical community for coverage led to the creation of the National Oncologic PET Registry (NOPR). Medicare coverage for selected cancer indications was provided only to PET facilities that participated in the registry [54]. Data supporting the hypothesis that use of PET changed patient management were weak [35]. Nonetheless, Medicare extended coverage for the initial diagnosis and staging of cancer. Both the lung volume reduction surgery and PET for cancer cases demonstrate the difficulty of rescinding coverage once it is offered provisionally through CED.

In 2006, CMS issued a CED policy allowing coverage for percutaneous transluminal angioplasty and stenting (PTAS) for prevention of a second stroke in high-risk Medicare beneficiaries only when they were enrolled in an FDA-approved trial. The results of this study, published in 2011, showed that patients undergoing PTAS have a much higher rate of stroke or death (14.7%) as compared to patients receiving medical management alone (5.8%) (Chimowitz 2011) [55]. As a result, enrollment of patients in the trial stopped earlier due to high risk of early stroke in patients undergoing PTAS and Medicare withdrew coverage for high-risk patients. The use of data for decision making for other Medicare CED initiatives has been hampered by lack of funding so that some studies never got underway, slower than anticipated trial enrollment, and opportunistic use of ongoing studies to provide evidence that ultimately did not provide they type of evidence Medicare needed to make informed decisions [17].

Apart from the Medicare program, there is one state-based program that is experimenting with CED. The one completed study was used to inform state coverage policy and was a study of spinal cord stimulation (SCS) for failed back surgery syndrome. This prospective, controlled cohort CED study was initiated by Washington State to understand the effectiveness and risks of SCS for chronic back and leg pain after spine surgery [56]. This study evaluated outcomes of workers’ compensation recipients with failed back surgery syndrome who received SCS with those who either: 1) received pain clinic evaluation with no SCS or 2) received neither SCS nor pain clinic evaluation. The SCS procedure was covered only for patients enrolled in the study [57]. After an assessment of both safety and efficacy of the treatment, there was no evidence for greater effectiveness of SCS as compared to the alternative treatments. At six months, SCS showed a small advantage in improving leg pain, only with higher use of opioids; and the effect disappeared in the long-term. Since this procedure was associated with no benefits beyond 6 months and entailed risks, including one life-threatening event, state policymakers continued to maintain non-coverage for spinal cord stimulators for treatment of failed back surgery syndrome.
Box 3: France
Which entities are involved in the process?

In France, the two entities involved in PBRSAs are the Transparency Commission (TC) and the Pricing Committee (PC). The TC gives an advice on access to reimbursement (SMR) and rates the relative effectiveness of drugs (ASMR), and it will often require post-launch observational study focusing on the use of a new product in real life. The results of such studies are also of interest for the PC, which will adjust its price-volume requirements to the results of the observational studies. The PC can, independently of the TC, also ask for specific studies.

Manufacturers must get prior approval from the TC for their planned design of the post-launch study. This advice is given by an internal expert group of the French HTA Agency (HAS), but does not commit the HAS.

At termination, the study results are evaluated by the Directorate for Evaluation of the HAS, which may require external advice. The results of the evaluation are transmitted to the TC and the PC if the latter is involved. At the time of reassessment, the TC can use other date alongside the post-launch study, and the company can also provide the TC with complementary data, so it is difficult to assess the specific impact of the post-launch study.

Public authorities sometimes claim that manufacturers play the clock and delay the launch of studies. In the new law on drugs, which is under legislative consideration, financial penalties will be increased for manufacturers which do not comply with the TC’s requirement in a timely way. Manufacturers sometimes complain that the studies requested by the TC are either not appropriate (observational studies can seldom lead to a controlled assessment of efficacy, for example), or unrealistic (use of a given molecule will change), or too broad (requiring a study of all available treatments in the same indication) and that delays in studies are mainly due to slow assessment of protocols by the internal expert group.

What is the general approach and experience in France?

The main types of products that are subject to such requirements are (a) expensive products, with high uncertainties at the time of launch, or (b) drugs for the treatment of high prevalence diseases, such as cardiovascular diseases and associated risk factors, diabetes, drugs for mental health, and antibiotics.

The results of such studies are to be used for the five-year reassessment of each drug by the TC, and also for the price readjustment. However, there may not be a pre-specified relation between the results of the study (performance) and reassessment. About 140 post-launch studies have been requested since 2000, but there is little public evidence of how their results have affected the reassessment, although such a study could be possible, comparing the first TC assessment to consecutive ones. But it is more difficult to assess the impact on prices since pricing agreements based on post-launch studies are kept confidential.

The products subject to such requirements are mainly expensive products, with high uncertainties at the time of launch, or drugs for the treatment of high prevalence diseases. In some cases, the TC will ask manufacturers with competing products to collaborate over a common registry, or cohort study. In a much publicized study on COX-2 inhibitors (CADEUS)—at the time of the launch of Vioxx and Celebre—the manufacturers involved were asked to fund a public research centre study to assess the claim of better gastro-intestinal safety of COX-2s relative to NSAIDs. Before the results of the study were published, Vioxx was withdrawn from the market, and Celebrex had its price cut.

Another typical example of such studies is when the public authorities anticipate that there may be off-label use of a drug, with high budget impact. In this case, the TC will ask for a study of actual use of the drug to identify use in off-label indications, and the pricing committee would eventually set a target of acceptable off-label use, which—if not me—would have an impact on price rebates.

The PC may be changing its position on PBRSAs. The previous PC chairperson argued that if there are too many uncertainties, manufacturers should invest more in their clinical development before they ask for reimbursement, and could expect to be rejected or get a low price to deal with the uncertainty. However, if the drug is promising, manufacturers will have to provide the payer with additional observational data, and reimbursement and prices will be renegotiated according to the outcomes of study. If manufacturers make a reasonable claim on an attribute of a new treatment that cannot be demonstrated in a trial, to ask for a premium price, if they can provide specific post-launch data to sustain this claim, then the PC can accept a higher price. The new PC chairperson has publicly declared an interest in PBRSAs, stating that a better price could be granted if outcomes of studies were positive. He expects such agreements to be as simple as possible, so as to have a rapid answer to questions of uncertainty raised at the time of launch.

Examples of PBRSAs in France

Amongst the 140 post-launch studies, two have been given some publicity, because they are specifically PBRSAs. In both cases, the Pricing Committee was the main instigator of the agreements.

One involves DPP4 inhibitors for patients with T2 diabetes. DPP4 inhibitors were a new class. Most of them have claimed reimbursement as an alternative to sulfamides in association with metformin, when first line metformin has failed. They have not demonstrated better efficacy for glycaemic control, but they present with a better tolerance profile. There is also experimental data to suggest that their efficacy to lower HbA1C lasts longer, delaying treatment escalation. At the first round of negotiation, the PC was ready to offer only a small premium price over existing alternatives for bi-therapy. The manufacturers were able to convince the PC to let them proceed to a large real life study to demonstrate their claim on durability. Thus, they were able to obtain a better price than expected, but with the condition that if the study did not support their claim, they would have to pay back the difference between the agreed-upon price and the initial price retrospectively for all sales. All new DPP4 inhibitors have been subject to the same agreement.

The second involves a controlled release form of risperidone, for the treatment of schizophrenic patient. In this case, the TC concluded that there was no major improvement versus the traditional one, and granted an ASMR 5 rating, leading to the same price as the existing presentation. The company argued for a better ranking, claiming that the controlled release form would lead to more patient compliance and so to fewer hospital admissions. The PC accepted the performance of a post-launch study to demonstrate a potential reduction of the number of hospital admissions, and accepted a higher initial price. The study’s results supported the company’s argument.

Box 4: Italy

Which entities are involved in the process?

In Italy, the two entities involved in PBRSAs are the Italian Medicines Agency (AIFA) and the National Health Service (NHS). In order to guarantee the affordability of innovative medicines, it is viewed as pivotal to have an approach able to link the use of medicine to the clinical outcomes obtained. The lack of evidence in the real-world clinical setting, in particular for innovative medicines, has motivated AIFA to use conditional reimbursement schemes (also known as Managed Entry Agreements) and Monitoring Registries to collect data on safety and effectiveness.

What is the general approach and experience in Italy?

The very first registries were implemented in 2006 in order to guarantee the following aims: control the overall budget; finance only cost-effective medicines; and to obtain additional data on clinical effectiveness. The AIFA Monitoring Registries are online tools. Patients’ Case Report Forms must be filled in using a specific web-based Monitoring Register. Since 2006, a total amount of 75 therapeutic indications related to 49 drugs were recorded in the Monitoring Registries. The largest number is for oncology, followed by diabetes, psoriasis, and those for orphan diseases. The Cancer Drugs Register covers all of the prescription centres in Italy with a population of over 100,000 oncology patients. All of the schemes involve payment for the first cycles and then full reimbursement by the NHS for responders and treatment discontinuation in non-responders. They differ as to whether a discount is given on the first cycles of treatment and whether a full or partial refund is given by the manufacturer for non-responders. AIFA uses the terms (a) “cost sharing” (when there is a price reduction for initial treatment cycles until it is clear whether a patient is responding), (b) “payment by results” (when the manufacturer reimburses the payer for non-responders) and (c) “risk sharing” (when only 50% of the costs of the non-responders are reimbursed by the manufacturer). Payment by results (with a full refund for non-responders) is the agreement form most commonly used.

Examples of PBRSAs in Italy

An example of a medicine for which a registry was created is Rasilez (aliskiren), used to treat essential hypertension. In Italy, Rasilez was reimbursed by the NHS only after the compilation of a registry by specialist centers. AIFA also included Rasilez in its Monitoring Register for Cardiovascular Medicines to investigate the safety profile. Two years of observational data showed that Rasilez reduced both systolic blood pressure and diastolic blood pressure in patients enrolled and had a good safety profile. The Registry was a very useful tool to examine the utilization of Rasilez as a 2nd line treatment. Based on the data presented in the Register the following decisions were taken: the shift of prescribing responsibility to GPs; a price reduction of Rasilez to align the price to the other hypertensive medicines; and introduction of a pharmaceutical expenditure ceiling. In order to define possible pharmaceutical expenditure ceiling, a budget impact analysis was completed, projecting the utilization of Rasilez as a 2nd line treatment.

An example of a “cost-sharing” type of scheme in oncology involves two oral drugs for renal cell carcinoma: sorafenib and sunitinib. In this scheme, the hospital receives a discount of 50% for the first 2/3 months of treatment. If a patient responds, the treatment is reimbursed and the discount is dropped [6].
Box 5: Value of Information and the Evaluation of PBRSAs

For any PBRSA scheme to be viable *ex ante*, but particularly for CED schemes (OWR and OIR), there must be the potential for value to be generated by further evidence generation, and the expected value of that information must exceed the expected cost of the scheme designed to generate that evidence. Despite this, the role of formal value of information techniques at the *ex post* evaluation phase of a scheme is limited. This is because we cannot, as a starting point for such an analysis, simply take the assessment of the expected value of perfect information (EVPI) before the scheme was implemented and net out the EVPI left after the evidence generation to estimate the value for the data that were collected. The reason for this is that EVPI describes the *expected* value of perfect information across all possible realisations of where uncertainty resolves. However, in any given realisation of a particular study, although parameter uncertainty will usually be reduced, consideration has to be given to the effect of that realisation on “decision uncertainty,” which also drives the EVPI calculation. It turns out that in some circumstances, the value of information following a PBRSA scheme may increase, not because of parameter uncertainty, but due to increased decision uncertainty.

For example, consider the first panel of Figure 2 which shows the estimated cost-effectiveness of a new technology as a distribution of incremental net-monetary-benefit (INMB) prior to a PBRSA. The potential cost-effectiveness is shown by a positive mean INMB, but with a wide variance and a 16% chance that the decision could be incorrect—the measure of decision uncertainty. Following a PBRSA, the uncertainty in the INMB is reduced as is indicated by the more precisely estimated INMB distribution in the second panel of the figure. However, because the location of the distribution is now closer to the decision threshold (INMB=0) (the mean expected INMB is less than the prior estimate), the decision uncertainty has increased to 28%. In the case illustrated, the EVPI after the PBRSA is approximately 50% greater than that which existed before the PBRSA. It should be noted, however, that the EVPI is calculated as the integration of the probability of an incorrect decision and the consequences (INMB loss) associated with that decision. In other words, whilst decision uncertainty looks at the probability of a wrong decision, the EVPI calculation weights that probability by the consequent loss. Thus, although Figure 2 illustrates a scenario where there is both an increase in decision uncertainty and in EVPI after the PBRSA, it is possible to show examples where decision uncertainty increases but the EVPI nevertheless decreases (because most wrong decisions lead to small losses) and vice versa.
Figure 2: Parameter and decision uncertainty prior to a PBRSA and after the scheme reports.
Box 6: Netherlands

Which entities are involved in the process?

In the Netherlands, the four primary entities involved in PBRSAs are the Dutch Healthcare Authority (NZa), the Healthcare Insurance Board (CVZ), the Medicinal Products Reimbursement Committee (CFH), and the Ministry of Health, Welfare, and Sport. Organizations (viz., industry associations of health care professionals) can formally submit a request for inclusion of a medicine in the policy regulation to the NZa. The NZa asks the CVZ for advice on the inclusion of a medicine in the policy regulations and CVZ bases its advice to the Ministry on an assessment of the medicine by the CFH.

What is the general approach and experience in the Netherlands?

In 2006, in the Netherlands, a coverage-with-evidence-development (CED) policy regulation was introduced for expensive medicines and orphan drugs. Costs for these drugs had to be covered under fixed hospital budgets, causing the phenomenon of ‘post code medicine’, exhibiting substantial geographic variation in usage. In order to remove the financial obstacles that hospitals encountered in using expensive medicines, the Minister of Health, Welfare, and Sport decided to adjust the policy for financing expensive intramural medicines and orphan drugs by introducing this type of PBRSAs. If a drug is listed, hospitals will be funded 80% of the drug expenses in addition to their budget, so only 20% will have to be covered from the hospital budget. Hospitals are fully reimbursed for orphan drug costs. Manufacturers, care-providers, and health care institutions will have to assure jointly that sufficient data are collected for re-assessing the effectiveness, cost-effectiveness, and actual indication the drug is used for, within a time frame of four years and for this purpose a study proposal should be submitted. This period was defined at three years at the very beginning, but this was experienced as being too short.

The CFH will assess the “therapeutic value” of the new drug in comparison with the standard or usual treatment (either medicinal or non-medicinal in nature) for a given indication in the Netherlands. Determination of therapeutic value is based on efficacy, cost-effectiveness, side effect profile, experience, applicability, and ease of use, although cost-effectiveness is not part of the advice CVZ provides over inclusion in the policy regulation. Once the therapeutic value is considered positive, a further criterion is that the prognosis for total expenses for the drug (based on a retail price that should be comparable to the prices for the same drug in neighboring countries) are equal or higher than 0.5% of the nationwide hospital expenses on medicines, which means approx. 2.5 million Euro expenses for the new drug. For orphan drugs the criteria is that total costs are equal or higher than 5% of the average costs of medicines in the academic hospitals only.

By the end of 2011, twenty-six drugs for thirty-four indications were on the positive list of this policy regulation, as well as ten orphan drugs. Thus far, five drugs have reached the re-assessment phase after four years. For two of these, no re-assessment dossier was submitted to CVZ and subsequently these drugs were taken off the positive list so the additional funding for the expenses for these drugs within hospitals stopped. The dossiers presenting the effectiveness, cost-effectiveness, and actual indication the drug is used for were not submitted to CVZ in 2011. Only one of the dossiers presenting real-world effectiveness, cost-effectiveness, and actual use was considered to be of adequate quality, while two of the three were assessed by CFH as inadequate. These findings were presented to the Insured Package Advisory Committee (ACP, part of CVZ), which still has to provide a final advice to the Minister of Health, Welfare, and Sport.

An example of a PBRSAs in the Netherlands

The first two drugs to be included (December 2006) on the positive list of the policy regulation were Remicade (infliximab) for the indication of psoriasis and Mabcampath (alemtuzumab) for advanced chronic lymphocytic leukemia. In 2008, Mabcampath was again put on the positive list for the indication of 1st and 2nd line treatment of chronic lymphocytic leukemia. This example shows that the reimbursement decision for drugs in the Netherlands is related to specific indications and the same drug can be assessed for inclusion in the expensive medicines regulation several times also, for example, Mabthera (rituximab) is on the list for rheumatoid arthritis, non-Hodgkin’s lymphoma, and lymphatic leukemia.
References


[34] Raftery J. Costly failure of risk sharing scheme. BMJ. 2010;340:1282-1284


[38] Barros PP. The simple economics of risk-sharing agreements between the NHS and the pharmaceutical industry. Health Econ. 2011; 20: 461-70.


[50] Lilford RJ. MS risk sharing scheme. Response from chair of scientific advisory committee. BMJ. 2010; 341: c3590.


