VALUE IN HEALTH 16 (2013) 703–719

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

ISPOR TASK FORCE REPORTS

Performance-Based Risk-Sharing Arrangements—Good Practices for Design, Implementation, and Evaluation: Report of the ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force

Louis P. Garrison Jr., PhD (co-chair)[1,*], *Adrian Towse, MA, MPhil (co-chair)*[2], *Andrew Briggs, MSc, DPhil*[3], *Gerard de Pouvourville, PhD*[4], *Jens Grueger, PhD*[5], *Penny E. Mohr, MA*[6], *J.L. (Hans) Severens, PhD*[7], *Paolo Siviero, BA*[8], *Miguel Sleeper, ACMA*[9]

[1]*Pharmaceutical Outcomes Research & Policy Program, Department of Pharmacy, University of Washington, Seattle, WA, USA;* [2]*Office of Health Economics, London, UK;* [3]*Institute of Health and Wellbeing, University of Glasgow, Glasgow, UK;* [4]*ESSEC Business School, Cergy Pontoise, France;* [5]*F. Hoffmann - La Roche AG, Basel, Switzerland;* [6]*Center for Medical Technology Policy, Baltimore, Maryland, USA;* [7]*Institute of Health Policy & Management, Erasmus University Rotterdam, Rotterdam, The Netherlands;* [8]*Agenzia Italiana del Farmaco, Rome, Italy;* [9]*Access to Medicines GlaxoSmithKline plc, Brentford, UK*

ABSTRACT

There is a significant and growing interest among both 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 amount or level of reimbursement is based on the health and cost outcomes achieved. There has always been considerable uncertainty at product launch about the 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 was 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 and recommendations include the following. 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. For good research practice in PBRSAs, it is critical to match the appropriate study and research design to the uncertainties being addressed. Good governance processes are also essential. The information generated as part of PBRSAs has public good aspects, bringing ethical and professional obligations, which need to be considered from a policy perspective. The societal desirability of a particular PBRSA is fundamentally an issue 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 and the accompanying reduction in uncertainty. The *ex post* evaluation of a PBRSA should, however, be a multidimensional exercise that assesses many aspects, including not only the impact on long-term cost-effectiveness and whether appropriate evidence was generated but also process indicators, such as whether and how the evidence was used in coverage or reimbursement decisions, whether budget and time were appropriate, and whether the governance arrangements worked well. There is an important gap in the literature of structured ex post evaluation of PBRSAs. As an innovation in and of themselves, PBRSAs should also be evaluated from a long-run societal perspective in terms of their impact on dynamic efficiency (eliciting the optimal amount of innovation).

Keywords: access with evidence development, conditional licensing, coverage with evidence development, managed entry schemes, outcomes-based, patient access schemes, pay for performance, performance-based risk-sharing arrangements, risk-sharing.

Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

* *Address correspondence to:* Louis P. Garrison Jr., Pharmaceutical Outcomes Research & Policy Program, Department of Pharmacy, Health Sciences Building, H375 1959 NE Pacific Street, H-375A, Box 357630, University of Washington, Seattle, WA 98195, USA.

E-mail: lgarrisn@u.washington.edu.

1098-3015/$36.00 – see front matter Copyright © 2013, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

http://dx.doi.org/10.1016/j.jval.2013.04.011

Background to the Task Force

Since 2007, there has been an acceleration in interest in a variety of arrangements between medical product manufacturers and payers that tie postlaunch data collection to payments. The ISPOR Performance-Based Risk-Sharing Arrangements Good Practices Task Force was approved by the ISPOR Board of Directors in March 2011 to set out the standards that should be applied to these arrangements, encompassing the design, implementation, and evaluation of such agreements. The report builds on previous work undertaken at Banff, in the UK Pharmaceutical Price Regulation Scheme, and by others as well as relevant work undertaken by other ISPOR Good Research Practice Task Forces, notably those tackling issues around the design, collection, and use of observational data to improve the quality of decision making.

Professors Lou Garrison and Adrian Towse, task force co-chairs, chose task force members on the basis of their knowledge and experience in decision modeling, study design, market access, coverage with evidence development, and performance-based pricing arrangements. Members represented a diverse range of practice and perspectives, including government (Agenzia Italiana del Farmaco [AIFA]), academia, health economic research and policy organizations, as well as the pharmaceutical industry. The task force was international with members from France, Italy, The Netherlands, Switzerland, the United Kingdom, and the United States.

The Task Force met approximately once a month by teleconference to develop and revise the outline and draft, as well as to discuss issues that arose in the process. A face-to-face meeting was held in November 2011 to develop recommendations and to reach consensus on content issues. In addition, the task force chairs had a series of one-on-one teleconferences to revise sections of the manuscript. All task force members reviewed and provided frequent feedback via oral or written comments on the manuscript drafts.

Preliminary findings were presented in a forum at the 2011 ISPOR 14th Annual European Congress in Madrid, Spain. Updated findings were presented at the Third Plenary Session of the ISPOR 17th Annual International Meeting in June 2012 in Washington, DC. In addition to the oral comments received during the two presentations, a draft of this article was distributed to the 100+ person ISPOR Performance-Based Risk-Sharing Arrangements Task Force Review Group in January 2012. During the Review Group round of comments and the final manuscript review sent to the entire ISPOR membership, several hundred written comments were received from 104 ISPOR members and organizations.

All comments, most of which were substantive and constructive, were considered. The comments were reviewed and discussed by the task force in a series of teleconferences, and addressed as appropriate in a revised manuscript. Once consensus was reached by all authors, the final report was submitted to *Value in Health* in April 2013.

All written comments, as well as the task force's responses, are published at the ISPOR Web site on the task force's Web page: http://www.ispor.org/Taskforces/performance-based-risk-sharing-arrangements.asp The task force report and Web page may also be accessed via the ISPOR homepage (www.ispor.org) via the purple Research Tools menu, Good Practices for Out comes Research. A list of reviewers is also available via the task force's Web page.

Introduction

There is a significant and growing interest among 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 study 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. (See [2] and [3] for comprehensive lists of PBRSA examples.) This broad trend across many developed countries 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 the 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 anticipated in the real world has increased along with the rising price of the new treatments, whether a biologic, device, or other medical technology. If payers are reluctant to adopt, manufacturers face the risk of reduced revenue for a product they regard as delivering value. PBRSAs represent one mechanism for reducing 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 terminology, a public good—one person's use of the information generally does not keep others from using it—regardless of whether it is generated by public or private entities. Public authorities who negotiate and fund evidence-generating arrangements need to follow good research practices (GRPs) to improve the quality of the information derived and to make the results of that research public where possible. Private insurers, who may have less legal obligation for transparency, can still benefit from GRPs 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, as long as it does not inappropriately deter them from agreeing to PBRSAs.

The objective of this Task Force report was 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

PBRSAs fall under a variety of names and categories: outcomes-based schemes, risk-sharing agreements, coverage with evidence development (CED), access with evidence development, patient access schemes (PASs), conditional licensing, and managed entry schemes [2,4–10]. For the purposes of this discussion, we group all these under the broad term "performance-based risk-sharing arrangements" (PBRSAs).

Box 1 – United Kingdom

Which entities are involved in the process?

The United Kingdom has "Patient Access Schemes" (PASs) defined by the Pharmaceutical Price Regulation Scheme (PPRS) of 2009. PASs are agreement-specific. Most are, however, "financial" arrangements intended to provide the UK National Health Service (NHS) with effective discounts from list price rather than being linked to health outcomes. The UK PASs include performance-linked reimbursement agreements and coverage with evidence development only with research schemes, but are mainly types of discount agreements.

What is the general approach and experience in the United Kingdom?

Examples of UK schemes include the following:

- The dose-capping agreement that the National Institute for Health and Care Excellence (NICE) entered into over ranibizumab (Lucentis) for macular degeneration could be seen as an effective price discount or a performance guarantee. Cost-effectiveness to NICE was acceptable only if the NHS paid for up to 14 injections per eye of eligible patients. Novartis will bear the costs of treatment beyond this [12].
- NICE recommended ustekinumab (Stelara) for severe plaque psoriasis on the condition that Janssen-Cilag ensures that the costs of treating patients weighing more than 100 kg will be no more than those of patients weighing less than 100 kg [13].
- The bortezomib (Velcade) agreement ensures the identification of responders. There is retrospective payer reimbursement for nonresponders. Responders receive further doses of the product at the normal price.
- Pazopanib (Votrient) for advanced kidney cancer involves a discount on the list price to achieve price parity with sunitinib (Sutent). GSK agreed additionally to give the NHS a financial rebate if pazopanib proves inferior to sunitinib with regard to its efficacy, in a head-to-head trial. The results of the COMPARZ (COMParing the efficacy, sAfety and toleRability of paZopanib vs. sunitinib) study were reported in October 2012 and showed noninferiority in progression-free survival.

In total, there are 28 schemes as of March 2013, of which 15 are simple discounts, 7 involve free stock, 2 involve dose capping, and 4 are more complex—these include the bortezomib and pazopanib schemes [14].

Evidence on the costs and effects of schemes is limited. Although the UK PASs are largely discount arrangements rather than performance-based risk-sharing arrangements (PBRSAs), the experience is relevant. Williamson [15] reports on a survey of oncology pharmacists in 31 NHS hospitals. Transaction costs for the NHS were the biggest concern. 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 nonresponders" [15]. In contrast, however, a review by the Department of Health (DH) focused on the additional numbers of patients receiving access to drugs deemed cost-effective by NICE (after including transaction costs) [16].

An example of a PBRSA in the United Kingdom

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-quality-adjusted life-year threshold. The UK MS RSS was negotiated in 2002 between the UK DH 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 3 years rather than the expected 18 months to recruit 5000 patients at 73 centers. The results of the 2-year assessment of accumulated disability of the 5000 patients recruited were not reported until 2009, 7 years after the agreement to have a scheme. In reporting the results, Boggild et al. [17] said that "the outcomes so far obtained in the prespecified primary analysis suggest a lack of delay in disease progression." Prices, however, were not adjusted downward 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) [18]; the usefulness of the Expanded Disability Status Scale as the outcome measure; and the impact on the choice of the comparator when evaluating subsequent new drugs for the same indications.

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

1. *There is a program of data collection* agreed to between the manufacturer (or the provider, in some instances) and the payer. It may be initiated or required by the payer—to address uncertainties about long-term effectiveness (beyond trial duration and 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. In some cases, the data collection is for patient group/population-based studies; in other cases, individual patients are tracked.
2. *This data collection is typically initiated during the time period following the regulatory approval* (which may be full, conditional, or adaptive), and linked to postlaunch coverage decisions. It is therefore directed at informing payers, providers, and prescribers as decision makers and is not intended as postregistration licensing requirements for further evidence.
3. *The price, reimbursement, and/or revenue for the product are linked to the outcome of this program of data collection* either *explicitly* by a preagreed rule or *implicitly* through an option to renegotiate coverage, price, or revenue at a later date. In some cases, reimbursement is linked directly to the performance of the drug in a particular patient—a form of individual performance guarantee.
4. *The data collection is intended to address uncertainty* about one or more of the following:
   - efficacy or effectiveness in the tested population as compared with the current standard of care;

- the efficacy or effectiveness in a broader, more heterogeneous population than used in registration trials or in prelicensing testing;
- the effects on long-term or more clinically significant end points than those included in registration trials (which—in the case of a drug—may have used surrogate markers) or in prelicensing studies (e.g., for procedures or devices);
- any adverse effects and adherence issues;
- 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 due to fewer hospital visits;
- the proportion of patients who will respond, that is, achieve a preset (minimum) outcome that may be an intermediate/surrogate end point;
- the numbers and types of patients likely in real-world practice to be treated with the new therapy; and
- whether the patients treated are the "right" ones, that is, they have attributes matching those patients whom, on the basis of current evidence, the payer is willing to fund (which may or may not include off-label use).

5. *These arrangements provide a different distribution of risk* between the payer and the manufacturer than does the historical manufacturer-payer relationship.

Arrangements that are simply disguised price discounts—and are not concerned with clinical performance—are excluded from this definition. UK PASs include a number of these [11]. However, a number of operational aspects that underpin a successful PBRSA apply to such schemes. (See Box 1 for more information on UK PBRSAs.)

From a broader societal perspective [19], a PBRSA can be thought of as an investment to gather more data to resolve one or more of the above-mentioned uncertainties. Any number of stakeholders may be involved in developing PBRSAs, including drug and device manufacturers, public and private payers and insurers, employers financing insurance, hospital and physician providers, central pricing authorities, and regional budget-holders. For purposes of this report, the discussion is generically framed in terms of manufacturers and payers as the two principal parties involved, recognizing that there can be provider-payer and provider-manufacturer arrangements as well as other variants. 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 or utilization than the payer thinks is justified given the evidence. The payer is concerned about "decision uncertainty"—the probability of paying for a product that might not be effective or cost-effective in some or all of the patients who receive it following adoption in their health care system [20].

Investment in a PBRSA should lead to an arrangement that will better 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, an analyst would 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 come from aligning incentives in a way that promotes optimal research and development. Indeed, only the former are usually considered explicitly and the latter are at best implicit.

Taxonomy for PBRSAs

Previous Taxonomies

As noted in the previous section, there are many terms that have been used to describe the types of arrangements considered in this article. It is helpful to distinguish those that fit our definition from those that do not. To draw these distinctions, we are able to build on previous taxonomies that have been published in the literature [2,8,21,22]. For example, McCabe et al. [8] recognize that such schemes do not always include a research component and

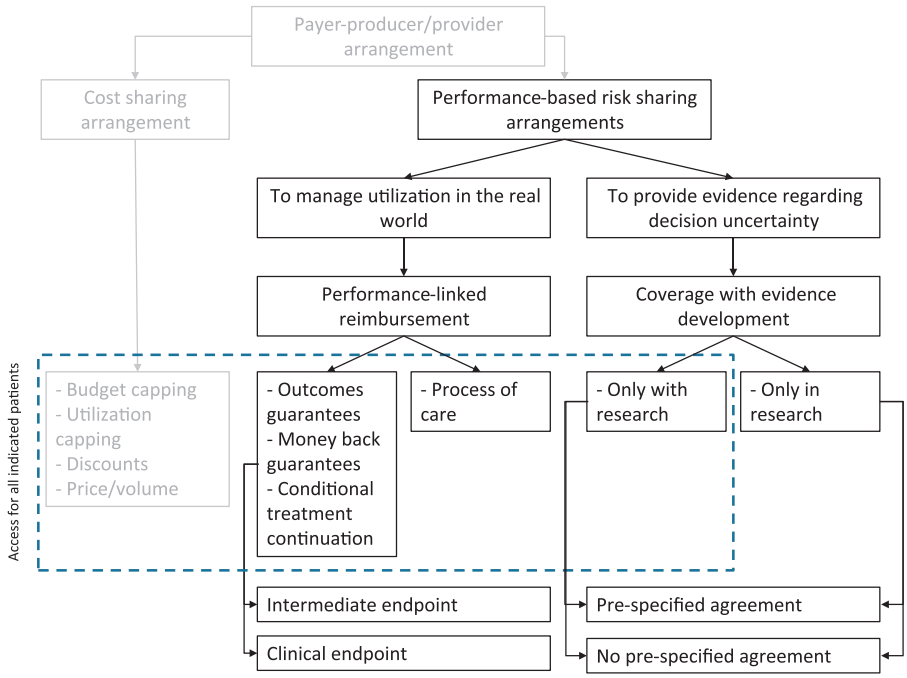

Fig. 1 – PBRSA Taxonomy. PBRSA, performance-based risk-sharing arrangement.

may focus on the health outcomes achieved for the individual patient, rather than at a population level.

Another taxonomy developed by Carlson et al. [2] was based on an inventory of published schemes categorized in terms of timing, execution, and health outcomes. This taxonomy made a clear distinction between schemes that are based on health outcomes and those that are not. Towse and Garrison [21] in their taxonomy also made the distinction between outcome-based and non–outcome-based and also between 1) those with agreements that specified how evidence would be translated into revisions to price, revenues, and/or use and 2) those that instead specify an evidence review point where renegotiation would occur . They made the point that the outcome evidence can come from a number of sources and study designs, including a randomized trial on a subset of patients (who may not necessarily be in the same health system) or an observational study of the patients being treated. They also drew a distinction between 1) uncertainty about the performance of the drug within a subgroup of patients and 2) uncertainty as to the subgroups of patients who would in practice receive the drug.

For purposes of defining GRP, we distinguish those payer-manufacturer arrangements that measure health outcomes in characterizing performance from those that focus on costs (non-health outcomes) (Fig. 1). We consider the latter "cost-sharing arrangements" only, thus falling 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 [23]. A case can be made that some of these agreements include an element of population health outcomes in that they are designed to cap payment at the point at which the population of users is equal to the target indicated population for which the payer regards the drug as good value. The Australian price-volume agreements have this rationale [21]. But given their focus on cost containment rather than measuring outcomes, GRPs are less relevant.

Among PBRSAs, we follow McCabe et al. [8] and further distinguish those that attempt to directly manage utilization and guarantee cost-effectiveness (i.e., "utilization-based") from those that include a strong research element (i.e., "research-based"). 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 cost-effective use of the technology. However, research-oriented arrangements are focused on covering the procedure for a period of time to develop further evidence that will reduce decision uncertainty about, for instance, the long-term outcomes expected to be achieved in groups of patients.

Current Taxonomy

Drawing on the previous taxonomies, Figure 1 also depicts these distinctions. From this figure, it can be seen that PBRSAs are a specific group of schemes among all possible payer-manufacturer arrangements. Pure "cost-sharing arrangements" shown in gray on the left-hand side are not within our definition of PBRSAs or scope for this report.

The taxonomy separates two types of PBRSA schemes:

- CED schemes whose goal is to provide coverage while the evidence is developed and
- performance-linked reimbursement schemes whose goal is to manage utilization, aiming to control the cost-effectiveness of a new technology in the real world through performance-linked reimbursement.

Box 2 – United States

What is the general approach and experience in the United States?

The process for deciding when to do a performance-based risk-sharing arrangement (PBRSA) in the United States 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 United States [26]. These initiatives largely focus on devices and surgical procedures [2]. 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 one half of US PBRSAs have been randomized controlled trials where access is granted only in research (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 only with research (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 the treatment of three diseases: metastatic breast cancer, epithelia carcinoma, and multiple myeloma [30]. NIH-funded trials demonstrated that the risks outweighed the benefits for the treatment of metastatic breast cancer, leading to the removal of coverage for that indication. FEHBP members, however, 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 to obtain coverage.

There have been few subsequent attempts at PBRSAs in the private sector in the United States. None of these has 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 [31]. Conditional coverage (OWR) was also offered by one payer in an agreement with a device manufacturer to examine the long-term durability of interventional procedures for the treatment of uterine fibroids [26].

The group most commonly associated with PBRSAs in the United States is the Medicare program, which initiated a CED program [32]. Medicare distinguishes between policies in which coverage is provided only to study participants (coverage under study protocol, i.e., OIR) and those that offer broad access but require

additional data collection (coverage with appropriateness determination, i.e., OWR). 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 15 CED policies; all but 2 have been coverage under study protocol, and most have targeted devices or diagnostics. Recent draft guidance would eliminate the OWR variant, enable the agency to extend CED to older, established technologies, and reiterate the ability for its regional administrative contractors to make individual assessments about the coverage of experimental technologies in research under the agency's clinical trials policies [33]. One regional Medicare administrative contractor has since issued its own policy on CED [34].

Apart from the Medicare program, there are two state-based programs that are experimenting with CED: Washington (with spinal surgery) [35] and Minnesota (with rehabilitative services for autism) [36].

Examples of PBRSAs

Notably, Medicare has used the evidence generated from CED policies to inform subsequent coverage determinations on 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 [37]. This well-designed NIH-funded trial found that apart from a small subpopulation, 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.

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. Medicare coverage for selected cancer indications was provided only to those PET facilities that participated in the registry [38]. Data supporting the hypothesis that the use of PET changed patient management were weak [28]. 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, the Centers for Medicare & Medicaid Services issued a CED policy allowing coverage for percutaneous transluminal angioplasty and stenting (PTAS) for the prevention of a second stroke in high-risk Medicare beneficiaries only when they were enrolled in a trial approved by the Food and Drug Administration (i.e., OIR). The results of this study, published in 2011, showed that patients undergoing PTAS have a much higher rate of stroke or death (14.7%) than do patients receiving medical management alone (5.8%) [39]. As a result, enrollment of patients in the trial was stopped earlier because of the 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, and so some studies never got underway or had slower than anticipated trial enrollment, and by opportunistic use of ongoing studies to provide evidence that was not the type of evidence Medicare needed to make informed decisions [26].

An illustration of a state-based program is the Washington State study used to inform state coverage policy for 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 [40]. 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 (i.e., OIR) [35]. After an assessment of both safety and efficacy of the treatment, there was no evidence for greater effectiveness of SCS as compared with the alternative treatments. At 6 months, SCS showed a small advantage in improving leg pain, but only with higher use of opioids, and the effect disappeared in the long term. Because this procedure was associated with no benefits beyond 6 months and entailed risks, including one life-threatening event, state policymakers continued to maintain noncoverage for SCS for failed back surgery syndrome.

CED PBRSAs

CED is a payer-manufacturer 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 within the target population who are "exposed" to the technology. A CED scheme with either a preagreed adjustment or a later renegotiation could lead to a change in use (i.e., in the number of subpopulations) and/or a price change. The result could be more efficient, sending manufacturers a clear signal about societal willingness to pay, as well as restricting or increasing use to the appropriate population(s).

CED PBRSAs can be further subdivided into two groups. PBRSAs in the first group apply to all new indicated patients who might be treated by using the new technology, with research taking place only in a subset or even in another health system (this is called "only with research"—OWR). PBRSAs in the second group are based only on those patients who were voluntarily included in an observational study or trial [24] (this is termed "only in research"—OIR). In the OIR case, not all patients who fall within the indication will have access or exposure to the new technology. We include both as examples of PBRSAs relying on CED. Walker et al. describe the criteria that might lead to the use of either an OWR or OIR scheme [25]. CED schemes can exist either with a prespecified agreement or without: for example, a public payer might be able to mandate them. PBRSAs without a prespecified agreement will require renegotiation (to adjust price, coverage, or revenues) at a later time based on the results of the research [26–29]. (See Box 2 for more information on CED in the United States.)

Performance-Linked Reimbursement PBRSAs

In contrast to these CED arrangements, the other category of PBRSAs is of those that aim to manage utilization to achieve cost-effective use 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 the process of care, and these PBRSAs can also be considered as programs to promote good quality of care. This means that reimbursement is specified ex ante to depend on the clinical decision-making process, for example, a provider's compliance with clinical guidelines or the selection of individual patients on the basis of a biomarker, such as a genetic test.

Other performance-linked schemes focus on ex post reimbursement, measuring intermediate or clinical end points. These arrangements include 1) "outcomes guarantees"—meaning payment for responders only—or 2) "conditional treatment continuation"—meaning payment for the continued use of the product based on intermediate end points. Thus, in contrast to CED, reimbursement in these schemes is generally determined for each individual patient, either prospectively by rule or retrospectively by result. Nevertheless, it would generally be possible, in theory, to aggregate the data collected in such real-world utilization to the population level to adjust overall payments post hoc. In addition, for research purposes, post hoc analyses to support additional population-level decision making are a possibility if the relevant patient-level data have been 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 together will have to address four research-related, good practice questions concerning the following: Q1) the desirability of the PBRSA (as opposed to some other form of reimbursement or research arrangement), Q2) the choice of research design, Q3) the approach to implementation, and Q4) the evaluation method to be used. In some instances, the payer and the manufacturer (or the 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, some or all of the four good practice questions will also be relevant, depending on the nature of the arrangement.

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

When Is a PBRSA Worthwhile or Desirable?

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 [21,41]:

1. Adopt (or partially adopt) despite the uncertainties, with the option to revisit the decision if more information becomes available.
2. Refuse to adopt until the manufacturer supplies better evidence to address the uncertainties.
3. Demand or mandate a lower price such that the uncertainty about value for the payer is reduced.
4. Enter into a PBRSA that a) manages utilization/outcome at the patient level or b) is a form of CED in which evidence is collected across patients for a review, potentially leading to prespecified adjustments or later ad hoc adjustments in price or utilization.

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—can be used for weighing these options [42-44]. The general desirability question is one as to whether a PBRSA can effectively and efficiently address the uncertainties that remain following marketing authorization.
Issues that would need to be considered include the following:

- the expected value(s) of research options in terms of reduced uncertainties for the parties to the agreement. The value of any additional research (and, therefore, of a CED or research-based PBRSA) will depend in part on the price of the medical product and its expected use in the patient population;
- the direct cost (in terms of out-of-pocket costs and administrative burden) of collecting evidence;
- the opportunity cost (in terms of expected health loss in the population) of any delays in access that might result from the use of a scheme; and
- 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 [3,25]. Where research may be made less feasible by adoption, then conducting research in another health system may enable research to take place alongside use (i.e., OWR) or, in other cases, the PBRSA can restrict adoption to a form of OIR.

A utilization-based PBRSA can change the 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 utilization-based PBRSAs will need to be assessed alongside the other options of 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 negotiate 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 [22]. But to assess complexity and cost, the manufacturer will also need to address questions Q2 to Q4—evidence collection, implementation, and evaluation—while paying attention to the perspectives of payers, prescribers, providers, and patients.

In these instances, whether to propose or accept a PBRSA will be a business decision for the manufacturer and a business and/or political decision for the payer. When either party proposes the use of a PBRSA, it should have established that a valid and efficient process of evidence collection following good professional practice is feasible, and it should be realistic regarding acceptable levels of uncertainty as well as the cost of data collection and the implementation challenges of the scheme. Neither party is per se subject to a discussion of GRPs, although a payer may indicate scientific research criteria that have to be met. In any case, the resulting data need to be robust enough to address the key uncertainties whether the scheme operates at an individual patient level (a form of performance-based reimbursement) or at a population-level (a form of CED).

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

Box 3 – France

Which entities are involved in the process?

In France, the two entities involved in performance-based risk-sharing arrangements (PBRSAs) are the Transparency Commission (TC) and the Pricing Committee (PC). The TC gives advice on access to reimbursement to the Service Médical Rendu (SMR; Provided Medical Service) and rates the relative effectiveness of drugs (Amélioration du Service Médical Rendu [AMSR; Improvement of Provided Medical Service]), and it will often require a postlaunch 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 postlaunch study. This advice is given by an internal expert group of the French HTA Agency (Haute Autorité de Santé [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 data alongside the postlaunch study, and the company can also provide the TC with complementary data, and so it is difficult to assess the specific impact of the postlaunch study.

Public authorities sometimes claim that manufacturers unnecessarily delay the launch of studies. In the new law on drugs, which is under legislative consideration, financial penalties will be increased for manufacturers who do not comply with the TC's requirement in a timely way. Manufacturers sometimes complain a) that the studies requested by the TC are not appropriate (observational studies that will 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 b) that the delays in studies are mainly due to the 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 and so with potential high budget impact, such as cardiovascular diseases and associated risk factors, diabetes, drugs for mental health, and antibiotics.

The results of such studies are used in the 5-year reassessment of each drug by the TC, and also for the price readjustment. There is often, however, not a prespecified relation between the results of the study (performance) and reassessment. About 140 postlaunch studies have been requested since 2000, but there is little public evidence of how their results have affected the reassessment, although such a study, comparing the first TC assessment to consecutive ones, is possible. But it is more difficult to assess the impact on prices because pricing agreements based on postlaunch studies are kept confidential.

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 cyclooxygenase-2 (COX-2) inhibitors called CADEUS (COX2 Inhibitors and Non Steroidal Anti Inflammatory Drugs: Description of Users)—at the time of the launch of rofecoxib (Vioxx) and celecoxib (Celebrex)—the manufacturers involved were asked to fund a public research center study to assess the claim of better gastrointestinal safety of COX-2s relative to nonsteroidal anti-inflammatory drugs. Before the results of the study were published, rofecoxib was withdrawn from the market, and celecoxib 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 the actual use of the drug to identify use in off-label indications, and the PC would eventually set a target of acceptable off-label use, which —if not met—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 should expect to be rejected or to receive 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 the 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, and if they can provide specific postlaunch 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

Among the 140 postlaunch studies, 3 have received some publicity because they are specifically PBRSAs. In each case, the PC was the main instigator of the agreements.

One involves dipeptidyl peptidase-4 (DPP4) inhibitors ("gliptins") for patients with type 2 diabetes. DPP4 inhibitors are a relatively new class of oral antidiabetic medications. They have sought reimbursement as an alternative to sulfonamides in association with metformin when first-line metformin has failed. They have not demonstrated better efficacy for glycemic control, but they provide a better tolerability profile. There are also experimental data to suggest that their efficacy to lower hemoglobin A_{1C} 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 dual therapy. The manufacturers were able to convince the PC to let them proceed to a large real-world 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 results for the first gliptins should be available in the first half of 2013.

A similar agreement was made for glitazones: the claim was also made that this class of antidiabetic drugs would delay escalation to insulin therapy. The goal was not reached, and prices were adjusted downward.

The third involves a controlled-release form of risperidone for the treatment of patients with schizo-

phrenia. In this case, the TC concluded that there was no major improvement versus conventional treatment, and granted an ASMR 5 rating, leading to the same price as those of existing medications. The company argued for a better ranking, claiming that the controlled release form would lead to more patient compliance and thus to fewer hospital admissions. The PC accepted a higher initial price subject to the performance of a postlaunch study to demonstrate a potential reduction in the number of hospital admissions. The study's results supported the company's argument.

Since October 2012, companies have had a mandate to submit a cost-effectiveness analysis (CEA) at first request for coverage for drugs considered as innovative (ASMR 1–3) and with an expected high budget impact. For such products, a CEA will also be required at the 5-year review, based on "real-life data," alongside other postlaunch evidence required on effectiveness and safety. In none of these cases do these requirements for supplementary evidence prespecify what is to be provided.

clinical development time prior to launch, this risks 1) still not demonstrating a sufficient benefit or improvement in effect, 2) reducing the period of exclusivity associated with patent protection, and 3) enabling a competitor to come into the market with similar or better results before the manufacturer has entered the market.

For payers, denying reimbursement on the basis of significant uncertainty runs the risk of depriving patients of the potential benefits of a new medical product. On the 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 financial consequences and criticism for any losses if the product is not cost-effective and/or provides little or no benefit to patients. It may also, in these circumstances, provide a poor benchmark as a standard of care for comparison with future treatments.

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 manufacturers than for payers. Yet, the use of a PBRSA (as opposed to a denial) may be socially optimal, as well as in the commercial interest of the manufacturer. (See Box 3 on the postlaunch studies in France.)

PBRSAs, if not well designed, may facilitate the pursuit of suboptimal strategies by either manufacturers or payers. For example, in a CED scheme with no prespecified agreement as to how price changes (prospectively or retrospectively) or as to how use changes in the light of study findings, both parties may engage in suboptimal behavior. Manufacturers might define their price target according to the anticipated risk of not meeting the target (i.e., go for the highest price possible if they think that the study will not support their claim) and delay evidence collection for as long as possible. Payers might—if they have substantial bargaining power—set their price offer low, that is, well below their reservation (highest) price and set a high target outcome for the study to increase the probability of a product failing to meet its target. They may refuse to increase the (low) price they have offered to pay even if the study shows that the outcomes exceed the target linked to the preset price. The payoffs to such behaviors can be reduced by prespecified agreements—for example, requiring retrospective rebates to discourage such manufacturer behavior or by requiring prices to rise when the target outcomes are exceeded to discourage such payer behavior.

Finally, it is important to remember that the payer is the agent of plan beneficiaries, who can also be patients: PBRSAs are efficient only if they make beneficiaries as a whole better off in terms of health outcomes and plan costs. Of course, equity and other social factors can come into play in addressing willingness to accept uncertainty, for example, with regard to orphan drugs.

Q2. Evidence collection: Which PBRSA research design is most appropriate to collect evidence that addresses the relevant uncertainties?
The answer will depend on the nature and type of the uncertainty that the PBRSA evidence collection is trying to address:

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

The preferred study design will 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 on previous GRPs for specific types of studies. Previous ISPOR Task Forces have defined GRPs in a number of relevant areas: modeling, nonrandomized studies of treatment effects, randomized controlled trials (RCTs), retrospective database analysis, and prospective observational studies [45–49]. 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

Table 1 – Factors affecting selection of randomized vs. observational design.

Factors favoring randomized design	Factors favoring observational design
• Relative efficacy remains a question of interest even after premarket studies • Feasible to randomize • Prognostic variables are unclear; most variation in outcomes is unexplained • Biologic process of disease is not well understood • Risk of getting the answer wrong is large (large impact on mortality, resource use, or treatment patterns) • Prognostic variables are unclear; most variation in outcomes is unexplained • Modest anticipated differences in effect size	• Relative effectiveness is of interest • Effect size is relatively large and/or selection bias can be reasonably controlled • Relatively rare and serious adverse events are the most important outcomes of interest • Major focus is on adherence or compliance with therapy • Interest in associations of outcomes by patient subgroups, observed practice patterns • Risk of getting the answer wrong is low • Cost or resource utilization is the main interest

Adapted from J Compar Effect Res 2012;1(3):281–92 with permission of Future Medicine Ltd.

active area of research, particularly in the field of comparative effectiveness research [50,51].

The number of general options for research design to collect data postlaunch on new products is fairly limited and reasonably well defined. These include the following:

- a traditional targeted RCT, focusing on efficacy;
- a larger pragmatic clinical trial, randomized but with less rigorous entry inclusion or exclusion criteria and minimal interference with usual follow-up care;
- a prospective observational study of patients without randomization including a comparator; and
- a hybrid design that includes the use of prospective observational cohorts and retrospective data to provide a comparator.

Each of these designs has strengths and weaknesses in the context of PBRSAs. Table 1, adapted from Gliklich et al. [50], summarizes the factors that favor either a randomized design or an observational design. What is important for a PBRSA is that the study is designed to answer the question at hand, that is, to address the specific uncertainty that most increases the likelihood of a bad decision. Historically, there is a general consensus that properly sized RCTs are the strongest method for determining a treatment effect. They can be criticized, however, for having limited generalizability across practice settings as well as for measuring efficacy rather than effectiveness. Large pragmatic clinical trials can be costly and difficult to manage, but they can offer a better estimate of real-world effectiveness. Prospective observational studies, for example, using data collection from electronic patient records or disease registries, can be useful for estimating real-world effectiveness as well as the relationship between surrogate end points and long-term outcomes. (See Box 4 for more information on PBRSAs in Italy.) These studies can be comparative by including a comparator or 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 end points and long-term outcomes. They also require adjustment for selection bias.

Under randomized design, there is a continuum from more explanatory to more pragmatic designs. For assessing comparative or relative 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, for example, linked to a patient registry or postmarketing surveillance cohort studies, could be feasible to monitor the effectiveness and safety of investigated medicines in clinical practice. After the period of the study observation, results could be modeled to investigate whether the continued use of a PBRSA is cost-effective.

Although most PBRSAs to date (with the exception of a number of those in the United States) 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. If it is not possible to conduct an RCT alongside use of the product by the payer, then the RCT could be conducted either 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 [52,53].

GRP, as Hutton et al. [5] have emphasized, requires outcome measures to be selected with care. They should be clear, measurable, objective, realistically achievable, and relevant. All parties

Box 4 – Italy

Which entities are involved in the process?

In Italy, the two entities involved in performance-based risk-sharing arrangements (PBRSAs) are the Italian Medicines Agency (Agenzia Italiana del Farmaco [AIFA]) and the National Health Service (NHS). To guarantee the affordability of innovative and expensive medicines, it is viewed as pivotal to have an approach that links the use of a 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 [MEAs]) and monitoring registries to collect data on safety and effectiveness. Within an MEA context, various instruments such as price-volume agreements, cost-sharing, budget cap, monitoring registries, payment by results, risk sharing, therapeutic plans, and "AIFA notes" are used to manage budget impact, uncertainty around clinical- and cost-effectiveness, and appropriate utilization of medicines.

What is the general approach and experience in Italy?

One of the most important instruments for MEAs in Italy is drug-monitoring registries. For example, the Cancer Drugs Register covers all the prescription centers in Italy with a population of over 100,000 oncology patients.

These registries aim to assess and track patient eligibility, evaluate utilization in clinical practice, collect epidemiological data including data on the safety profile, and collect additional information that was missing at the first evaluation stage. This should guarantee appropriate use of medicines according to their therapeutic indications.

The AIFA monitoring registries are online tools. Patient case report forms must be filled in by using a specific Web-based monitoring register. Since 2006, a total of 78 therapeutic indications related to 66 active compounds were recorded in the monitoring registries, distributed as follows: 30 for antineoplastic drugs, 14 for orphan drugs, 1 project for the treatment of psoriasis, 1 for a cardiovascular drug, 2 for ophthalmic drugs, 2 for rheumatoid arthritis drugs, 2 for diabetes drugs, 2 for dermatological drugs, 2 for respiratory drugs, 1 for an osteoporosis drug, and 2 specific projects for multiple sclerosis and attention deficit/hyperactivity disorder . Of the 78 therapeutic indications using a monitoring registry, 28 are also part of a conditional reimbursement agreement. AIFA uses the terms "cost sharing" (when there is a price reduction for initial treatment cycles until it is clear whether a patient is responding), "payment by results" (when the manufacturer reimburses the payer for nonresponders), and "risk sharing" (when only 50% of the costs of the nonresponders are reimbursed by the manufacturer). Cost sharing applies a general discount to all eligible patients at the beginning of treatment, whereas both risk sharing and payment by result use a payback mechanism to compensate for the treatment costs of nonresponders. In terms of implementation, the system of applying an initial discount to all eligible patients used in the cost-sharing scheme is simpler to administer than the system of reimbursement for nonresponders used in the risk-sharing and payback schemes.

In terms of the mix of schemes, there are 12 indications reimbursed under a cost-sharing scheme, 2 indications under a risk-sharing scheme, and 14 indications under a payment-by-results scheme. All three

models include a health outcomes element in the form of evaluation of the treatment efficacy and continuation of treatment conditional on a positive response to the drug. The main differences lie in the financial arrangements.

Examples of PBRSAs in Italy

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

An example of a cost-sharing 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 [2].

will need to ask of the choice of outcome measure and the study design:

- Is the type of evidence being promised sufficient to address the uncertainty in both quantitative and qualitative terms?
- Are the end points 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 outcomes for PBRSAs will also be influenced by the scope of regulatory approval, especially for drugs. If the manufacturer wishes to use a PBRSA's results for promotion, any claims need to be consistent with the label.

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 [54], and by Hutton et al. [5]. 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 health care system arising from the PBRSA should be proportionate to the potential gains. For example, we note the requirements expressed in the UK Pharmaceutical Price Regulation Scheme that relate to both CED and performance-linked reimbursement schemes [11]:
  - When considering the burden, the full cost—in terms of both direct monetary costs and the opportunity cost of staff time—to the health care system of any PAS 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 system 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 et al. [5] suggest that a PBRSA study period longer than 3 years faces the risk of becoming increasingly irrelevant in the face of changing clinical practice and technological advancement. 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 the 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 this may not always be the case. There are examples where governments have agreed to fund arrangements (e.g., Medical Services Advisory Committee in Australia and Catalan Agency for Health Technology Assessment in Spain). The United Kingdom Multiple Sclerosis Risk Sharing Scheme (UK MS RSS) was jointly funded by the Department of Health and the companies involved [17,55].
- *How is responsibility for undertaking data collection and analysis allocated?* In a CED scheme, it should be subject to normal research governance arrangements as described in the next section.
- *Is the data collection efficient?* For efficiency purposes and to answer questions about effectiveness, the design will often make use of extensive data collected routinely (e.g., claims data, lab and pharmacy data, and 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.
- *What will be the process for reviewing and analyzing 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" [5]. In some agreements, there will be a preagreed link between the evidence and the final decision (e.g., an adjustment to price to ensure a cost-per-quality-adjusted life-year of a specific amount) that will be subject to 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 (e.g., based on provisional results)?* In performance-linked reimbursement schemes, such as responder-based reimbursement, this is integral to the scheme provided that such provisional arrangements support the needs of the

health care organization and avoid being excessively burdensome.

Governance

Some PBRSAs are bilateral commercial agreements between a private payer and a manufacturer, such as the case between Merck and Cigna for selected antidiabetic drugs in the United States [31]. In these situations, a formal governance structure is not essential. This is also the case when the arrangements are between a public payer and a manufacturer for utilization schemes (as in Italy [Box 4] or the United Kingdom [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 only-with-research patient access scheme for pazopanib [Votrient]).

In other cases, PBRSAs involve agreements among multiple stakeholders, and the need for formal governance structures is greater. For example, in the United States, the creation of the implantable cardiac defibrillator 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 [26]. The implantable cardiac defibrillator 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 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 [56].

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 a channel for receiving broad public comment.

In these circumstances, it is essential that there be a formal governance structure in place to ensure transparency of the nature and aims of the scheme, accountability, and a means to mitigate conflicts. 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 [55,57].

Transparency of process is important. A more controversial issue is transparency about any CED pricing arrangement. The value of new products—and therefore their price—will vary among different payers and settings, and revealing prices may lead to some payers referencing lower prices elsewhere rather than looking to pay for the value established by the scheme in their own health system. Thus, transparency of price is unlikely to contribute to the efficiency of a PBRSA.

Effective governance will require the following:

- The governance committee should have a clear charter specifying who is involved and their respective roles. Sign-off procedures for research design and study protocols are recommended between the governance committee and principal investigators.
- 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 upfront with clear information about the types of information or products each funder will receive as a result of its 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 and to achieve the clearly stated initial aims of the scheme, it is desirable to have a process in place for the independent review of research designs and the neutral, independent conduct of the research.

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; for example, they have to report whether an intervention "worked or not." From an evaluation standpoint, however, 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 could be a public good, there would seem to be a case for some explicit requirements for the disclosure of such public information. When a private payer and manufacturer agree to a PBRSA, 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.

In recent years, however, with the advent of public reporting requirements, such as clinical trials.gov, there seems to be a growing recognition that when these activities involve the voluntary participation of patients, there is an ethical and professional obligation to report on research results—at least for prospective clinical trials.

The public goods aspects of the information generated by PBRSAs should not be overlooked. The incentives for manufacturers, public payers, and private payers to rely on research evidence generated by others are always there [4,58–60]. From a global (societal) perspective, free riding on evidence already generated is a good thing in the short run but it reduces the incentives for investment in new evidence collection by any individual payer and/or manufacturer if the appropriation of evidence affects its competitive position. Even where payers and/or manufacturers are not competing, they will underinvest because they are not aware of the value to others of the evidence that they might generate through a PBRSA. New PBRSAs that follow GRPs may need to be subsidized or encouraged to 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 including the following: 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?

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. Some national health systems (and some private plans) have the stated or

implied aim of maximizing the health and related benefits of the covered population given a fixed annual budget. Any appraisal of a PBRSA in those schemes should consider the opportunity cost of health care resources in terms of health and related gains foregone.

Some of the additional costs of data collection, including negotiation, monitoring, and assessment, can be measured relatively easily. It will be more difficult to measure other costs, such as the foregone benefits to patients not receiving access during the data collection if the design is based on an OIR scheme. A full assessment will also need to consider reversal effects, that is, 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 behavior 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 performance-linked reimbursement scheme or a CED scheme) may have been developed with the principles of VOI 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 VOI generated for a single CED scheme. (See Box 5 for a discussion of this issue.)

Evaluation Questions

Because one cannot assess the VOI generated by a single CED scheme directly post hoc, there is a need to rely on process indicators of the scheme's success. It will be an important part of the design of any PBRSA (performance-linked reimbursement or CED) scheme to define the metrics by which the success of that scheme can be assessed. As per the above discussion, an evidence development scheme should unambiguously reduce *parameter* uncertainty: therefore, improving the precision of parameter estimates is indicative of success. Process indicators of success of the scheme should relate to the first three research questions, including the following:

- 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 and time?
- Was the integrity of the design/estimation maintained?
- Did the governance arrangements work well?
- Did the process to underpin a decision with further evidence prove successful?

The last point is very important. Ultimately, of course, to meet the objectives of a CED scheme, it is necessary to show that the decision making following the reporting of the scheme was informed by the additional evidence. Where an OWR scheme was implemented with a further review, 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, the question is 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? Where a performance-linked reimbursement scheme is being run, appropriate decision making will require the ability to show that the agreed outcome adjustments were made to improve the cost-effectiveness of the intervention.

Box 5 – Value of information and the evaluation of performance-based risk-sharing arrangements (PBRSAs)

For any PBRSA scheme to be viable ex ante, but particularly for coverage with evidence development schemes (both only with research and only in research), 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 (VOI) techniques at the ex post evaluation phase of a scheme is limited. This is because one should not, 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 realizations of where uncertainty resolves. However, in any *given* realization of a particular study, although parameter uncertainty will usually be reduced, consideration has to be given to the effect of that realization on "decision uncertainty," which also drives the EVPI calculation. In some circumstances, the VOI following a PBRSA scheme may increase, not because of parameter uncertainty, but because of 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, while 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 in which there is an increase in both decision uncertainty and in EVPI after the PBRSA, it is possible to show examples in which decision uncertainty increases but the EVPI nevertheless decreases (because most wrong decisions lead to small losses) and vice versa.

Evaluation Experience

Evidence on PBRSAs to date is limited:

- Puig-Peiró et al. [41] conducted a systematic literature review to identify existing knowledge about the costs and benefits, assessed either quantitatively or qualitatively, of PBRSAs. They found that more than 40% of the publications referred to the UK MS RSS, and no studies were able to evaluate the overall economic impact of a PBRSA. All studies only included

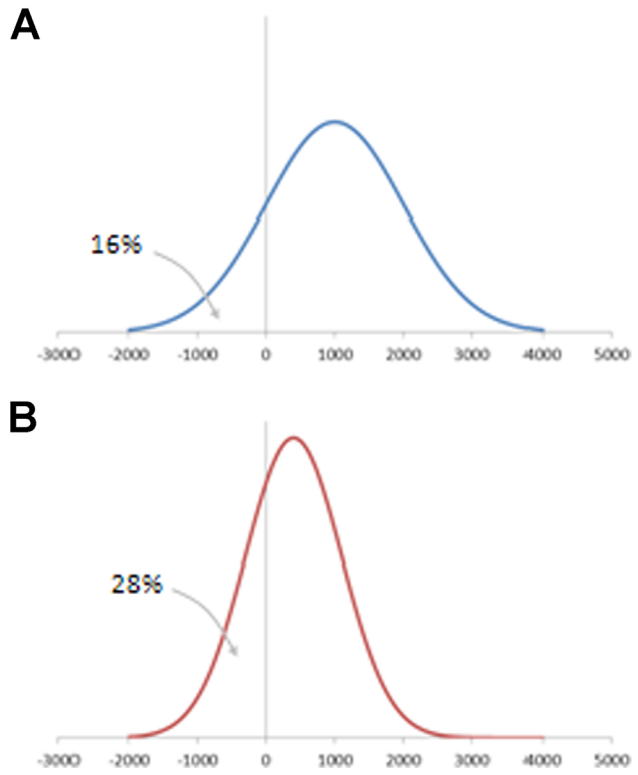

Fig. 2 – Parameter and decision uncertainty prior to a PBRSA (A) and after the scheme reports (B). PBRSA, performance-based risk-sharing arrangement.

qualitative discussions of costs and benefits, with the exception of the UK MS RSS in which some costs estimates were reported.

- In the case of the UK MS RSS, the results of the 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, and problems in the governance of the scheme [61].
- Neumann et al. [62] reviewed five PBRSAs in the United States and the United Kingdom and concluded that they are hard to implement in practice.
- The Italian PBRSA utilization 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).
- PBRSA experience in The Netherlands to date suggests that in some cases the quality of evidence collected has been poor (see Box 6).
- The experience of CED in the United States is covered in Box 2. The main form has 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.
- The available evidence on another non-US CED OWR scheme —bosentan (Tracleer) scheme in Australia—suggests that it was successful. Patients were enrolled in a registry. The results suggested that some reduction in price was appropriate. A competing product was listed at a 15% price discount, and the sponsor agreed to a government request for the same discount [63]. It is clear, however, that the interpretation of the results was much complicated than anticipated, leading to a range of possible price reductions [63].

Box 6 – The Netherlands

What entities are involved in the process?

In 2006, in The Netherlands, a coverage with evidence development (CED) policy was introduced for expensive medicines (i.e., total expenses for the drug account for 0.5% of the nationwide hospital expenses on medicines, which means 2.5 million euros) and orphan drugs. Organizations (viz., industry or associations of health care professionals) could formally submit a request for the inclusion of a medicine in the policy regulation to the Dutch Healthcare Authority (Nederlandse Zorgautoriteit [NZa]). The NZa asked the Healthcare Insurance Board (College voor zorgverzekeringen [CVZ]) for advice on the inclusion of a medicine in the policy regulations. CVZ based its advice to the Ministry of Health, Welfare, and Sport (Volksgezondheid, Welzijn en Spor [VWS]) on an assessment of the medicine by the Medicinal Products Reimbursement Committee (College Farmaceutische Hulp [CFH]) (from 2013 Wetenschappelijke Adviesraad [WAR]) and/or the Insured Package Advisory Committee (Adviescommissie Pakket [ACP]). CVZ assessed the "therapeutic value" of the new drug in comparison with the standard or usual treatment (either medicinal or nonmedicinal in nature) for a given indication in The Netherlands.

What is the general approach and experience in The Netherlands?

Determination of therapeutic value is based on efficacy, cost-effectiveness, side-effect profile, experience, applicability, and ease of use. In addition, the CVZ assessed the applicant's first estimate of potential cost-effectiveness together with a study proposal for the collection of real-world data for the reassessment of effectiveness, cost-effectiveness, and efficient use. If the drug is included under the policy, this study proposal with suggestions from the CVZ provides the starting point for the outcomes research. Manufacturers, providers, and health care institutions had to ensure jointly that sufficient data would be collected within 3 years for reassessing outcomes; however, the evaluation period was changed to 4 years, after 3 years was found to be too short.

By the end of 2011, 26 expensive drugs for 34 indications were on the positive list of this policy regulation, as well as 10 orphan drugs. By that time, five drugs had reached the reassessment phase after 4 years. For two of these, no reassessment dossier was submitted to the CVZ, and subsequently these drugs were taken off the positive list and the additional funding for these drugs within hospitals was stopped. For the other three drugs, the dossiers were submitted to the CVZ in 2011, presenting the effectiveness, cost-effectiveness, and actual use based on 4-year findings. None of the dossiers presenting real-world effectiveness, cost-effectiveness, and actual use was considered to be of adequate quality. These findings were presented to the ACP committee (part of the CVZ), which assessed the societal aspects of the interventions. To date, the CVZ has advised positively on only one of these products to the VWS, more specifically, a performance-linked scheme for omalizumab (Xolair) under which the drug would be reimbursed only if the effectiveness for an individual patient was demonstrated after a fixed period. Omalizumab is indicated for patients having severe persistent asthma. Because of the relatively high price, the cost-effectiveness was considered unfavorable. Based on an initiative of the manufacturer (Novartis), the Dutch society for pulmonologists (Nederlandse Vereniging van Artsen voor Longziekten en

Tuberculose [NVALT]), and the Longfonds (Lung Fund) representing patients, the VWS and the CVZ agreed to an arrangement where the manufacturer would reimburse the drug acquisition costs in cases in which a patient does not respond to omalizumab. This is defined in a general agreement, and the manufacturer will sign individual contracts with the separate hospitals. This scheme will be evaluated after 2 years.

Examples of PBRSAs in The Netherlands

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

Recently, ranibizumab (Lucentis) for the treatment of neovascular age-related macular degeneration was reassessed by the CVZ. It was proposed to the VWS to exclude ranibizumab from reimbursement beginning 2015 unless additional evidence shows that patients who do not respond or are contraindicated for off-label use of bevacizimab (Avastin) have a clinical benefit compared with current treatment options. To date, the ministry has not indicated whether it will follow this advice of the CVZ.

New 2012 Regulation

Beginning in 2012, the policy for expensive medicines and orphan drugs was terminated and replaced by a CED scheme for all health care technologies that fall within the reimbursement package (called "voorwaardelijke toelating"). Previously, CED for expensive drugs was intended to supplement the hospital budget. From 2012 on, CED is meant to cover interventions that are not yet part of the basic package due to lack of evidence on effectiveness. Conditional reimbursement explicitly aims to gather evidence on effectiveness. Based on this regulation, a restricted number of technologies can be temporarily included in the national reimbursement package for a period of 4 years. A restriction on numbers is considered necessary because the budgets and research capacity to gather the required evidence are both limited. Besides temporary inclusion in the reimbursement package, drugs can be temporarily funded (while excluded from the package) on the basis of a positive treatment effect without substantial evidence on cost-effectiveness and appropriate use. The criteria for choosing specific technologies for the Dutch CED system are mainly based on: the definition of the essential evidence gap; the specific research question and quality of a detailed research protocol; the feasibility to collect relevant evidence during the 4-year period; and the balance of the value of the evidence in relation to the research budget needed. VWS decides which interventions will be temporarily reimbursed, depending on the available research budget.

- Experience in the Asia-Pacific is dominated by Australia, which has had numerous confidential price-volume "managed entry" agreements with volume based on expected cost-effective utilization. At least four PBRSAs have been implemented in Australia: only bosentan has been reviewed in the published literature. In several other Asia-Pacific countries—notably, China, Taiwan, South Korea, and Japan—a few conditional treatment continuation PBRSAs have been initiated and others are under active consideration but there is no evaluation evidence [64].
- Although the UK PASs include discount arrangements as well as PBRSA performance-linked reimbursement and CED OWR schemes, the experience is relevant. Williamson [15] reports on a survey of oncology pharmacists in National Health Service hospitals. Transaction costs for the National Health Service were the biggest concern. However, in contrast, a review by the Department of Health focused on the additional numbers of patients receiving access to drugs deemed cost-effective by the National Institute for Health and Clinical Excellence (after including transaction costs) [16].
- Sweden has implemented 17 PBRSAs, of which 16 were CED, though no formal evaluations have been published [1].
- The results from other European Union countries are unclear, and the schemes are in evolution [14,65–68]. In a recent survey of oncology schemes in Europe, six European countries were found to be using either financially based schemes or outcome-based schemes, though most were financially based. But countries were found to vary in how they implement them. Opinions on arguments for and limitations of such schemes were surveyed, but there was no formal evaluation [68].

Overall, the literature suggests that there is an important gap in structured ex post evaluations of PBRSAs. Performance-linked reimbursement schemes appear to have been more successful to date than CED schemes. The evidence, however, is limited, mixed, qualitative, and partial.

Conclusions

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 GRPs 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 as follows:

- 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 use performance-linked reimbursement 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 (i.e., "OWR") or product use only in the context of research (i.e., "OIR").
- PBRSAs using CED can use observational studies or RCTs in either an OWR setting or an OIR setting. Where an RCT is preferred but may not be feasible with OWR, evidence can be collected from an RCT in another jurisdiction, provided the results can be translated into the setting of interest for stakeholders.
- All PBRSAs, including those tracking patients in performance-linked reimbursement schemes, can provide valuable evidence that is potentially 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 the 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. PBRSAs should include a prespecified deliberative process setting out when and how future decisions are to be tied to the additional evidence that is developed. Good governance processes are also essential.
- Because they can generate evidence that 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 use 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 expected 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.
- There is an important gap in the literature of structured ex post evaluation of PBRSAs. Performance-linked reimbursement schemes appear to have been more successful than CED schemes. The evidence, however, is limited, mixed, qualitative, and partial.
- The ability to run successful PBRSA schemes using either performance-linked reimbursement or CED will provide an important additional tool for increasing efficiency in health care. Robust evaluation of such schemes will be important for learning.
- As an innovation in and of themselves, PBRSAs should ultimately be evaluated from a long-run societal perspective in terms of their impact on dynamic efficiency (eliciting the optimal amount of innovation).

Acknowledgments

We thank the more than 100 ISPOR reviewers for their helpful comments: their names can be found along with our responses on the ISPOR Web site. Their feedback was very helpful in focusing the manuscript and our thinking, and it contributed to the consensus nature of an ISPOR task force report. We appreciate their time, expertise, and insights. We also thank Elizabeth Molsen of ISPOR for her hard work and dedicated support to our Task Force throughout the preparation of this report.

Source of financial support: No financial support was received.

REFERENCES

[1] Carlson JJ, Gries K, Sullivan SD, Garrison L. PHP146 Current status and trends in performance-based schemes between health care payers and manufacturers. Value Health 2011;14:A359–60.
[2] Carlson JJ, Sullivan SD, Garrison LP, et al. Linking payment to health outcomes: a taxonomy and examination of performance-based reimbursement schemes between healthcare payers and manufacturers. Health Policy 2010;96:179–90.
[3] Walker S, Schulpher M, Claxton K, et al. Coverage with evidence development, only in research, risk sharing or patient access scheme? A framework for coverage decisions. CHE Research Paper 77, University of York, April 2012.
[4] de Pouvourville G. Risk-sharing agreements for innovative drugs: a new solution to old problems? Eur J Health Econ 2006;7:155–7.
[5] Hutton J, Trueman P, Henshall C. Coverage with evidence development: an examination of conceptual and policy issues. Int J Technol Assess Health Care 2007;23:425–32.
[6] Garber AM, McClellan MB. Satisfaction guaranteed—"payment by results" for biologic agents. N Engl J Med 2007;357:1575–7.
[7] Cook JP, Vernon JA, Manning R. Pharmaceutical risk-sharing agreements. Pharmacoeconomics 2008;26:551–6.
[8] McCabe CJ, Stafinski T, Edlin R, Menon D. Access with evidence development schemes: a framework for description and evaluation. Pharmacoeconomics 2010;28:143–52.
[9] Department of Health and Ageing, Australian Government. Framework for the introduction of a Managed Entry Scheme for submissions to the Pharmaceutical Benefits Advisory Committee. Available from: http://www.pbs.gov.au/info/publication/factsheets/shared/framework-for-introduction-of-managed-entry-scheme-for-PBAC-submissions. [Accessed May 2, 2012].
[10] Klemp M, Fronsdal KB, Facey K. What principles should govern the use of managed entry agreements? Int J Technol Assess Health Care 2011;27:77–83.
[11] Towse A. Value-based pricing, research and development, and patient access schemes: will the United Kingdom get it right or wrong? Br J Clin Pharmacol 2010;70:360–6.
[12] National Institute for Health and Clinical Excellence. NICE Guidance TA155 Ranibizumab and Pegaptanib for Treatment of Age-Related Macular Degeneration. Manchester, UK: National Institute for Health and Clinical Excellence, 2008.
[13] SCRIP. NICE set to recommend Stelara for psoriasis. SCRIP World Pharmaceutical News 17 August 2009. Available from: http://www.scripintelligence.com/policyregulation/NICE-set-to-recommend-Stelara-for-psoriasis-173593. [Accessed 2012].
[14] National Institute for Health and Care Excellence. List of technologies with approved Patient Access Schemes, recommended by NICE for use in the NHS. Available from: http://www.nice.org.uk/aboutnice/howwework/paslu/ListOfPatientAccessSchemesApprovedAsPartOfANICEAppraisal.jsp. [Accessed March 15, 2013].
[15] Williamson S. Patient access schemes for high-cost cancer medicines. Lancet Oncol 2010;11:111–2.
[16] UK Department of Health. Patient access schemes and flexible pricing: statement of outcomes from the review of the pricing flexibility measures under Chapter 6 of the 2009 PPRS. Available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/146829/dh_132794.pdf. [Accessed February 23, 2012].
[17] Boggild M, Palace J, Barton P, et al. Multiple sclerosis risk sharing scheme: two year results of clinical cohort study with historical comparator. BMJ 2009;339:b4677.
[18] Lilford RJ. MS risk sharing scheme: response from chair of Scientific Advisory Committee. BMJ 2010;341:c3590.
[19] Garrison LP, Mansley EC, Abbott TA, et al. Good research practices in cost-effectiveness analyses: a societal perspective: The ISPOR Drug Cost Task Force Report—part II. Value Health 2010;13:8–13.
[20] Griffin SC, Claxton KP, Palmer SJ, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. Health Econ 2011;20:212–24.
[21] Towse A, Garrison LP Jr. Can't get no satisfaction? Will pay for performance help? Toward an economic framework for understanding performance-based risk-sharing agreements for innovative medical products. Pharmacoeconomics 2010;28:93–102.
[22] Stafinski T, McCabe CJ, Menon D. Funding the unfundable: mechanisms for managing uncertainty in decisions on the introduction of new and innovative technologies into healthcare systems. Pharmacoeconomics 2010;28:113–42.
[23] Zaric GS, O'Brien BJ. Analysis of a pharmaceutical risk sharing agreement based on the purchaser's total budget. Health Econ 2005;14:793–803.
[24] Chalkidou K, Hoy A, Littlejohns P. Making a decision to wait for more evidence: when the National Institute for Health and Clinical Excellence recommends a technology only in the context of research. J R Soc Med 2007;100:453–60.
[25] Walker S, Sculpher M, Claxton K, Palmer S. Coverage with evidence development, only in research, risk sharing, or patient access scheme? A framework for coverage decisions. Value Health 2012;15:570–9.
[26] Mohr PE, Tunis SR. Access with evidence development: the US experience. Pharmacoeconomics 2010;28:153–62.
[27] Briggs A, Ritchie K, Fenwick E, et al. Access with evidence development in the UK: past experience, current initiatives and future potential. Pharmacoeconomics 2010;28:163–70.

[28] Levin L, Goeree R, Levine M, et al. Coverage with evidence development: the Ontario experience. Int J Technol Assess Health Care 2011;27:159–68.
[29] Trueman P, Grainger DL, Downs KE. Coverage with evidence development: applications and issues. Int J Technol Assess Health Care 2010;26:79–85.
[30] Rettig R, Jacobson P, Farquhar CM, et al. False Hope: Bone Marrow Transplantation for Breast Cancer. New York: Oxford University Press, 2007.
[31] Carlson JJ, Garrison LP Jr., Sullivan SD. Paying for outcomes: innovative coverage and reimbursement schemes for pharmaceuticals. J Manag Care Pharm 2009;15:683–7.
[32] Centers for Medicare & Medicaid Services. Guidance for the Public, Industry, and CMS Staff: National Coverage Determinations with Data Collection as a Condition of Coverage: Coverage with Evidence Development. Baltimore, MD: Centers for Medicare and Medicaid Services, 2006.
[33] Centers for Medicare & Medicaid Services. Draft guidance for the public, industry, and CMS staff: coverage with evidence development in the context of coverage decisions. November 29, 2012. Available from: http://www.cms.gov/medicare-coverage-database/details/medicare-coverage-document-details.aspx?MCDId=23. [Accessed March 21, 2013].
[34] Palmetto GBA. Jurisdiction 1 Part B. MolDx: coverage with evidence development. Available from: http://www.palmettogba.com/palmetto/providers.nsf/DocsCat/Providers~Jurisdiction%201%20Part%20B~Browse%20by%20Topic~MolDx~General~8ZLQQ74482?open&navmenu=ll. [Accessed March 21, 2013].
[35] Turner JA, Hollingworth W, Comstock B, Deyo RA. Comparative effectiveness research and policy: experiences conducting a coverage with evidence development study of a therapeutic device. Med Care 2010;48(Suppl.):S129–36.
[36] Garrett E, Schiff JS, eds. Autism Spectrum Disorders: Report to the Minnesota Commissioner of Human Services. St. Paul: Minnesota Department of Human Services, 2013.
[37] Ramsey SD, Sullivan SD. Evidence, economics, and emphysema: Medicare's long journey with lung volume reduction surgery. Health Aff (Millwood) 2005;24:55–66.
[38] Hillner BE, Liu D, Coleman RE, et al. The National Oncologic PET Registry (NOPR): design and analysis plan. J Nucl Med 2007;48:1901–8.
[39] Chimowitz MI, Lynn MJ, Derdeyn CP, et al. [SAMMPRIS Trial Investigators]. Stenting versus aggressive medical therapy for intracranial arterial stenosis. N Engl J Med 2011;365:993–1003.
[40] Hollingworth W, Turner JA, Welton NJ, et al. Costs and cost-effectiveness of spinal cord stimulation (SCS) for failed back surgery syndrome: an observational study in a workers' compensation population. Spine (Phila Pa 1976) 2011;36:2076–83.
[41] Puig-Peiró R, Mestre-Ferrandiz J, Sussez J, et al. Literature review on patient access schemes, flexible pricing schemes and risk-sharing agreements for medicines. Value Health 2011;14: A233–510, RS1.
[42] Eckermann S, Willan AR. The option value of delay in health technology assessment. Med Decis Making 2008;28:300–5.
[43] Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics 2006;24:1055–68.
[44] Claxton K, Palmer S, Longworth L, et al. Informing a decision framework for when NICE should recommend the use of health technologies only in the context of an appropriately designed programme of evidence development. Health Technol Assess 2012;16:1–323.
[45] Berger ML, Dreyer N, Anderson F, et al. Prospective observational studies to assess comparative effectiveness: the ISPOR Good Research Practices Task Force Report. Value Health 2012;15:217–30.
[46] Berger ML, Mamdani M, Atkins D, Johnson ML. Good research practices for comparative effectiveness research: defining, reporting and interpreting nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—part I. Value Health 2009;12:1044–52.
[47] Johnson ML, Crown W, Martin BC, et al. Good research practices for comparative effectiveness research: analytic methods to improve causal inference from nonrandomized studies of treatment effects using secondary data sources: the ISPOR Good Research Practices for Retrospective Database Analysis Task Force Report—part III. Value Health 2009;12:1062–73.
[48] Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health 2005;8:521–33.
[49] Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—modeling studies. Value Health 2003;6:9–17.
[50] Gliklich RE, Leavy MB, Velentgas P, et al. Incorporating stakeholder perspectives in developing a translation table framework for comparative effectiveness research. J Compar Effect Res 2012;1(3): 281–92.
[51] Hefland M, Berg A, Flum D, et al. Draft Methodology Report: "Our Questions, Our Decisions: Standards for Patient-Centred Outcomes Research." PCORI Methodology Committee, Washington, D.C., 2012.
[52] Eckermann S, Willan AR. Globally optimal trial design for local decision making. Health Econ 2009;18:203–16.
[53] Eckermann S, Willan AR. Optimal global value of information trials: better aligning manufacturer and decision maker interests and enabling feasible risk sharing. Pharmacoeconomics 2013;18:203–16.
[54] Menon D, McCabe CJ, Stafinski T, Edlin R. Principles of design of access with evidence development approaches: a consensus statement from the Banff Summit. Pharmacoeconomics 2010;28:109–11.
[55] Pickin M, Cooper CL, Chater T, et al. The multiple sclerosis risk sharing scheme monitoring study—early results and lessons for the future. BMC Neurol 2009;9:1.
[56] Raftery J. Costly failure of risk sharing scheme. BMJ 2010;340:1282–4.
[57] Levine MN, Julian JA. Registries that show efficacy: good, but not good enough. J Clin Oncol 2008;26:5316–9.
[58] Antonanzas F, Juarez-Castello C, Rodriguez-Ibeas R. Should health authorities offer risk-sharing contracts to pharmaceutical firms? A theoretical approach. Health Econ Policy Law 2011;6:391–403.
[59] Barros PP. Pharmaceutical policies in European countries. Adv Health Econ Health Serv Res 2010;22:3–27.
[60] Barros PP. The simple economics of risk-sharing agreements between the NHS and the pharmaceutical industry. Health Econ 2011;20:461–70.
[61] Towse A, Garrison L, Puig-Peiro R. The use of pay for performance for drugs: can it improve incentives for innovation? Occasional Paper 11/05. Office of Health Economics, 2012.
[62] Neumann PJ, Chambers JD, Simon F, Meckley LM. Risk-sharing arrangements that link payment for drugs to health outcomes are proving hard to implement. Health Aff (Millwood) 2011;30:2329–37.
[63] Wlodarczyk J, Reid CM, Pater G. Funding linked to ongoing research: impact of the bosentan patient registry on pricing in Australia. Value Health 2011;14:961–3.
[64] Coulton L, Annemans L, Carter R, et al. Outcomes-based risk-sharing schemes: is there a potential role in the Asia-Pacific markets? Health Outcomes Med 2012;3:e205–19.
[65] Adamski J, Godman B, Ofierska-Sujkowska G, et al. Risk sharing arrangements for pharmaceuticals: potential considerations and recommendations for European payers. BMC Health Serv Res 2010;10:153.
[66] Garattini L, Casadei G. Risk sharing agreements: what lessons from Italy? Int J Technol Assess Health Care 2011;27:169–72.
[67] Lexchin J. Coverage with evidence development for pharmaceuticals: a policy in evolution? Int J Health Serv 2011;41:337–54.
[68] Espin J, Rovira J, Garcia L. Experiences and impact of European risk-sharing schemes focusing on oncology medicines. Eminet January 2011;1–39.