OHE Research
Office of Health Economics

Research Paper 12/03

Can and Should Value Based Pricing Be Applied to Molecular Diagnostics?

Martina Garau, Adrian Towse, Louis Garrison,

Laura Housman and Diego Ossa

April 2012

Can and Should Value Based Pricing Be Applied to Molecular Diagnostics?

Martina Garau[1], Adrian Towse[1], Louis Garrison[2], Laura Housman[3], Diego Ossa[3]

[1]Office of Health Economics, [2]University of Washington

[3]Novartis Molecular Diagnostics

OHE Research Paper 12/03

April 2012

For further information, please contact
Martina Garau
Senior Economist
Office of Health Economics
105 Victoria St
Tel: +44 (0)20 7747 8867
mgarau@ohe.org

About OHE Research Papers

OHE Research Papers are intended to provide information on and encourage discussion about a topic in advance of formal publication. Any views expressed are those of the authors and do not necessarily reflect the views or approval of OHE, its Editorial Board, or its sponsors. Once a version of the Research Paper's content is published in a peer-reviewed journal, this supersedes the Research Paper and readers are invited to cite the published version in preference to the original version.

Acknowledgement

The work on this paper was funded by Novartis Molecular Diagnostics.

Table of Contents

Abstract

Current pricing and reimbursement systems for diagnostics are not efficient. Prices for diagnostics often are driven by administrative practice and expected production cost. The purpose of this paper is to discuss how a value based pricing (VBP) framework that is being used to ensure efficient use and price of medicines also could be applied to diagnostics. Diagnostics not only facilitate health gain and cost savings, but also provide information to inform patients' decisions on interventions and their future "behaviours". For value assessment processes, we recommend a two-part approach. Companion diagnostics introduced at the launch of the drug should be assessed through new drug assessment processes considering a broad range of value elements and a balanced analysis of diagnostic impacts. A separate diagnostic-dedicated committee using VBP principles should review other diagnostics lying outside the companion diagnostics-and-drug "at-launch" situation.

1. Introduction

The UK is introducing value based pricing (VBP) for new drugs, building on the role of the National Institute for Health and Clinical Excellence (NICE) in assessing the cost-effectiveness of NHS treatments to improve the efficiency with which drugs are priced and used. Other countries also are using health technology assessment (HTA) to inform pricing and reimbursement decisions for medicines.

The VBP proposal from the Department of Health in the UK indicates that other factors not related to the quality-adjusted life year (QALY) should be considered, including: burden of disease (defined as combined unmet need and severity), the degree of therapeutic innovation, and health related benefits to patients not measured by the QALY. A societal perspective also should be taken, including benefits and costs outside health gain and health system costs (DH, 2011).

We can abstract beyond the specific UK proposals to identify the following elements of value.

1. Health effect is usually the single most important benefit of health technologies. Direct health effects can be measured using indicators of efficacy or effectiveness such as the QALY, which combines changes in the quality and length of life of an intervention.
2. Any cost-offsets within the healthcare system are a second key benefit. Savings to the health care system (offset by the additional cost of using the technology) are usually included in standard cost -effectiveness analyses.

Other elements of value fall into three distinct types.

1. A QALY's "value" to society may be higher or lower depending on who gets it. This might depend on the characteristics of the patients receiving the health gain (for example, age), on the nature of the illness in question, or on the pre-treatment level of health or disability of the patients (Shah, 2009). The UK VBP proposals suggest that the value of the QALY should be weighted by disease severity.
2. Elements of benefit to the patient that are not necessarily captured in the QALY (or any other measure of health gain), including:

a. Health related quality of life aspects not well reflected in a generic measure. For example, vitality is an important aspects of cancer patients' health, but it is not explicitly included in EQ-5D (Garau et al., 2011), which is one of the most used health measurement systems.
b. Health care process related aspects, such as being treated with dignity, at a convenient time and location, and after only a short wait. These may have health consequences, but the preference for them (as reflected in patients' stated preferences, or in political targets; for example, waiting times) goes beyond any health gain.

3. Information for the patient that, for example, enables life style choices to be made independent of any health effects
4. Other costs and benefits beyond those to patients and the NHS, such as the benefits to employers of getting people back to work more quickly and quality of life improvements for carers

The purpose of the paper is to discuss how a VBP framework also could be applied to diagnostics. In particular, we show how the incremental benefits generated by adding diagnostics to a health care pathway can include not only health gains (some of which may go beyond those captured by the QALY) and treatment cost savings, but also increased information available to patients to make decisions on treatment and/or their future lifestyle "behaviours".

We begin by outlining a framework with five pathways to identify the value of diagnostics. We then discuss three key process issues: aggregating value elements to inform price decisions, separating the value of test-treatment combinations, and designing optimal institutional processes for the value assessment of diagnostics. We conclude with policy recommendations.

2. Framework to Assess the Value of Diagnostics

Diagnostic tests include a broad range of techniques varying in their (1) level of complexity (from a simple clinical assessment to complex in vitro diagnostics assays) and (2) purpose (to determine the risk of developing a disease, the presence of a disease, an individual's prognosis, or treatment response). From an economic perspective, any type of diagnostic test can enhance the level of

information about a specific clinical condition or health state and so reduce or eliminate uncertainty (Garrison and Austin, 2007). If testing is linked to treatment or can improve disease management, then it generates downstream health effects of extended life and/or improved quality of life.

The value dimensions discussed below represent benefits that would be missed if a test were not available. We compare a situation where a test can be used *ex ante* to select the optimal intervention (including drug treatment and other interventions such as prevention) to a situation where a test is not available, in which case the most appropriate intervention is selected on a trial-and-error basis requiring *ex post* observation. Our framework applies specifically to molecular diagnostics[1], including companion diagnostics used to predict patients' responses to drug treatments (personalised medicine), tests to predict loss of treatment response (preventing the onset of severe stages of disease) and tests to predict disease risk.

The use of molecular diagnostics can generate value through the following five pathways[2].

1. Reducing or avoiding the adverse effects associated with treatment (including the medical and non-medical costs of managing them). Depending on the severity of the treatment side effects, testing can:
    a. allow a treatment to receive marketing authorisation by improving the benefit- risk ratio associated with the treatment
    b. increase adoption of the treatment, in cases where a treatment is licensed, but is not widely used because of its perceived unfavourable average benefit- risk balance when considered across a broad patient population

2. Reducing or avoiding time delays in selecting the most appropriate intervention. This has three main consequences:
    a. it generates health gain. When a disease is at an advanced stage (e.g. metastatic cancer), identifying non-responders and switching them to an alternative dosage,

[1] Molecular diagnostics are defined here as tests enabling molecular analysis of genes, proteins or metabolites.

[2] These five pathways are developed from Danzon and Towse (2002), which identifies reduced adverse reactions and targeted effects, and Garrison and Austin (2007), which identifies reductions in uncertainty. They are independent and additive.

treatment or care at the right time may have a significant impact on the patient's length and/or quality of life

b. it generates cost savings as it can avoid or reduce the cost of treating non-responders, including the cost of the drug[1]
c. it avoids or reduces inconvenience to patients who do not need to experience a long diagnostic process or try different therapies to identify the one most suitable

3. Increasing patient adherence or willingness to undertake preventative measures including changes in behaviour. Patients are more motivated if they know the intervention is likely to work. In the case of companion diagnostics, however, patients found to be non-responders might experience disutility as they can feel 'left-behind', and lose hope and even motivation to pursue any other, less effective, but appropriate therapy.

4. Enabling a treatment effective only in a small fraction of the population to be made available. This could happen by:
    a. "rescuing" treatments that may otherwise either not have been licensed or withdrawn because of the limited treatment effect across the overall population (i.e. favourable clinical effects in a subgroup are overwhelmed by the large group of non-responders)
    b. increasing the chance of a treatment meeting reimbursement criteria (if a diagnostic targeting responders improves cost -effectiveness), or being included in clinical guidelines (if evidence provided is deemed sufficient to change treatment protocols)
    c. accelerate the R&D process for treatments when a biomarker or other genetic characteristic allowing for patient stratification is ascertained at an early stage of treatment development. For example, patient stratification in oncology clinical trials could reduce attrition rates in overall clinical development and, in particular, attrition rates from Phase II to Phase III (Walker and Newell, 2009).

5. Reducing uncertainty about the value of potential new treatments and likely effectiveness of available treatments. In the first case, a test can improve information on the prevalence of a

[1] Strictly, from a societal perspective the relevant cost is the marginal cost.

particular untreatable condition, which in turn could help direct R&D toward that unmet need. In the case of available treatments, one type of uncertainty is around expected health effects and costs. This relates to knowledge of the disease (diagnosis, prognosis, casual explanation) and to the clinical and cost- effectiveness of treatments. It influences the risk of poor value for money for payers, i.e. the likelihood that treatments are not cost effective.

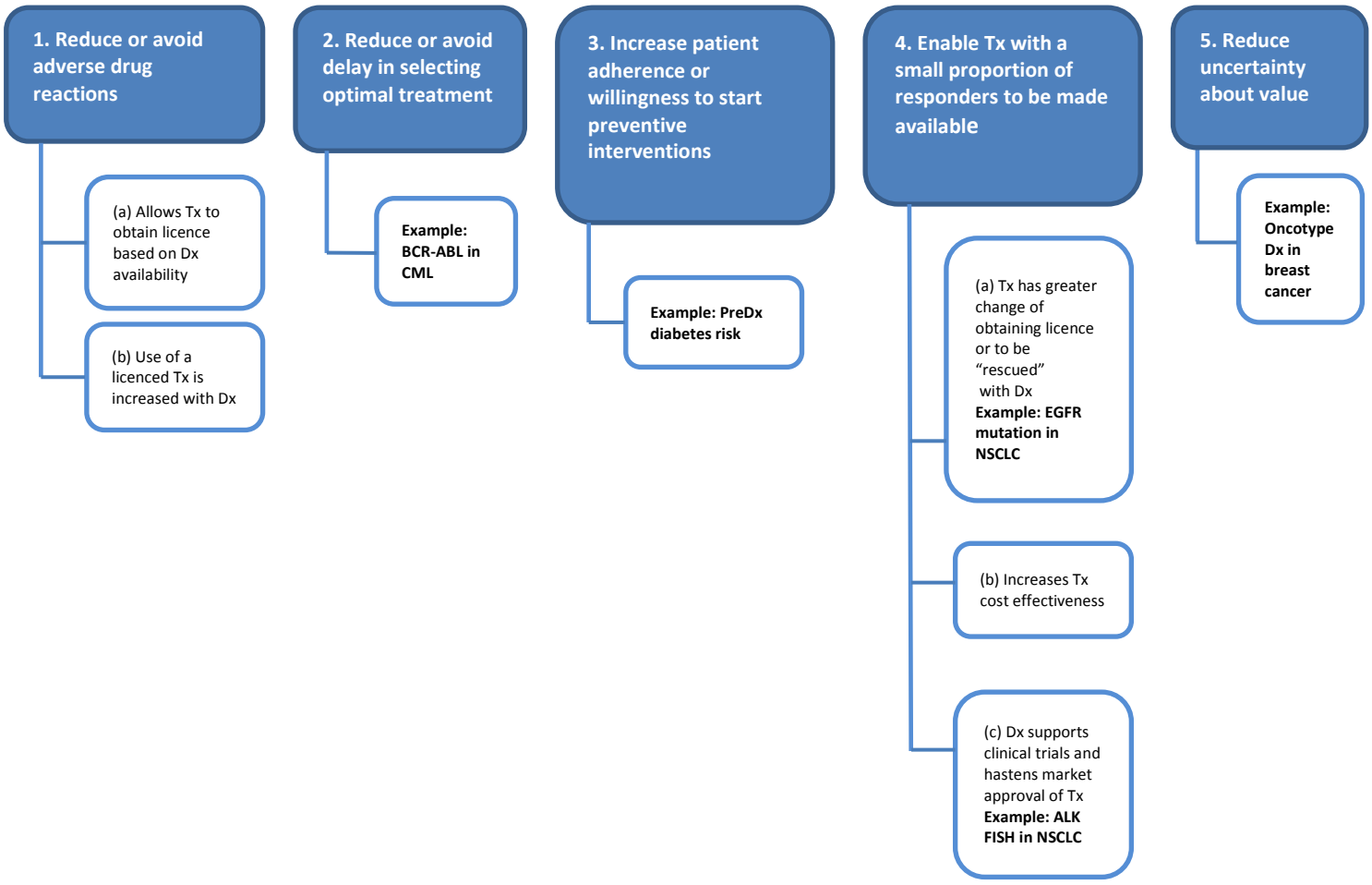
Another type of uncertainty is the perceived value of information to patients of reduced uncertainty as to their medical condition *independent of the expected health outcomes* (Han, Klein and Arora, 2011). The literature has defined it as the "value of knowing" or, as Ash, Patton and Hershey (1990) put it, "knowing for the sake of knowing", even if the condition is untreatable (Neumann et al., 2012). Patients may value information from a test regardless of the impact on their treatment strategy for the following reasons:

a. Decreased level of "ambiguity", a situation where probabilities of certain outcomes are highly uncertain (Ellsberg, 1961). There is evidence that people dislike ambiguous situations and prefer to receive information regardless of its nature ("bad" and "good" news) (Kenen, 1996; Neumann et al., 2012). In some cases, of course, test results can yield disutility. Patients are not always indifferent to the outcome of the test (Ash, Patton and Hershey, 1990). In the case of degenerative diseases such as Alzheimer's disease with limited treatment options and a high level of emotional burden, fear of living with the possibility of developing the disease can be very distressing. If the disutility associated with "bad news" is higher than the utility gain of "good news", then testing may not be to the benefit of the patient. In those situations, the choice as whether to test should ultimately be left to the individual patient as part of an informed decision making process involving all interested parties.
b. Provide reassurance to patients (value of "rule out"), particularly to those already identified as "at risk". A person with family history of a certain genetic disease can value a predictive test providing proof of the absence of, or lower chance of contracting, the disease in the future (Kenen, 1996).

In addition, individuals might want to undertake a test with no treatment options because the results will affect their family/life planning, including choices related to personal finances, work and leisure time (Lee, Neumann and Rizzo, 2010).

For these reasons, diagnostics have an impact on patients' utility beyond QALY-oriented health outcomes. These effects are not necessarily reflected in payers' decisions, which are usually based on health related quality of life measures not reflecting those aspects. Figure 1 matches these five pathways of value with illustrative examples of recently approved molecular diagnostic tests. Box 1 provides more details on those examples.

Figure 1: Pathways of value of molecular diagnostics and key examples



Dx = diagnostic; Tx = treatment; See Box 1 below for other abbreviations

Box 1: Case study examples

1. HLA-B*5701 is an allele associated with hypersensitivity to abacavir, which is part of a multi-drug regimen for HIV-1. This treatment was marketed before discovering the association between the HLA-B*5701 allele and the adverse reactions. Identification of the marker has increased prescribing of abacavir, which now is recommended for HLA-B*5701-negative patients in European and US guidelines.
2. The BCR-ABL test measures minimal residual disease of chronic myelogenous leukemia (CML). It can identify CML patients who are receiving treatment but not responding to it. The use of the BCR-ABL test can generate health gains as it can prevent the disease to progress to blast crisis and death. It also can lead to cost savings as it enables first-line treatment to stop when it no longer is effective.
3. PreDx Diabetes Risk test estimates the patient risk for developing Type 2 diabetes over the next five years. It was found to be better than most current methods of predicting Type II diabetes risk. This can further encourage patients to follow a healthy lifestyle and take other preventive measures.
4.
    a. Epidermal growth factor receptor (EGFR) mutation test predicts response to tyrosine kinase inhibitor (TKIs) treatments, such as gefitinib for non-small-cell lung cancer (NSCLC). Gefitinib initially was approved based on positive Phase II trial results, but subsequently withdrawn when Phase III failed to show a survival benefit. After the identification of EGFR mutations and its association positive response rate to TKIs, gefitinib received regulatory approval in the EU and other markets in combination with the EGFR mutation test.
    b. The human epidermal growth factor receptor 2 (HER2/neu) test is used in breast cancer to predict a patient's response to trastuzumab. For example, NICE in the UK recommends trastuzumab for advanced breast cancer and for adjuvant treatment of early-stage HER2/neu positive breast cancer as estimates of cost per QALY of the test-treatment combination were found below the standard threshold.
    c. The ALK FISH test is used in combination with crizotinib. The treatment recently licensed in the US targets a small subset -- between 3% and 8% -- of NSCLC patients with an ALK-positive molecular abnormality. Research on crizotinib started before the discovery that a fusion of two genes (ALK and EML4) could cause some lung cancers. However, the subsequent development of the ALK FISH test has accelerated the development process and increased the likelihood of crizotinib delivering health benefits and commercial value (Young, 2011).
5. Oncotype DX and MammaPrint are multi-gene assays that can identify patients with a high risk of recurrence, guide intervention decisions, and reduce the risk of dispensing unnecessary chemotherapy. Criteria currently used to predict risk of recurrence in breast cancer patients following surgery are not very accurate. As a result, many patients are either over- or under-treated with adjuvant chemotherapy.

The magnitude of the value created also will vary according to a number of test related parameters. In particular, low specificity and low sensitivity test accuracy will decrease the potential net gains to patients and to the health system from testing. For example, patients who are wrongly identified as responders (false positive) and those who are wrongly identified as non-responders (false negative) will miss the opportunity to receive a clinical decision from which they can benefit.

There are further implications if a diagnostic does not provide a binary response (positive or negative, yes or no). With a binary test, the overall patient population can be split into two subgroups –i.e. expected responders and non-responders. When a test does not provide a binary answer, for example, when it is aimed at measuring the level of individual protein expression, then there will be a patient subpopulation for which the test does not provide a clear-cut response. This subgroup potentially could benefit from the treatment, but uncertainty about treatment effectiveness is not substantially reduced with the use of a test. When this subpopulation is large relatively to the other two subsets (the “yes” and “no”) and the cost of the test is high compared to the cost of the treatment, the test-treatment combination might not be as cost-effective as the use of treatment on its own.

3. Options for Aggregating Elements of Value into a VBP

We have explored how diagnostics create value as well as incurring costs. Translating these elements of value into a value based price the payer is willing to reimburse requires the different types of "value" to be aggregated. The principal options for combining value elements that are not specific to processes for diagnostics are:

1. Converting all value into monetary terms – usually called the "net monetary benefit" approach (Claxton and Posnett, 1996)
2. Considering each type of benefit in terms of its own "unit of measurement"', and applying a set of weights to each benefit type to represent the rates at which different types of benefit may be traded-off with each other, and scores to indicate how well each benefit type is achieved by the medicine in question. This is called a multi-criteria decision analysis (MCDA) approach (Devlin and Sussex, 2011).
3. Selecting one principal measure of benefit, the default option being QALYs, as the "numeraire" and then up-rating or down-rating that measure using a series of weights to reflect the magnitudes of other types of benefit. Another option would be to assess (using stated-preference approaches) how people trade off QALY gains with other value elements such as informational benefits that are independent of health gains.
4. Using a "deliberative process" of the sort used by NICE and other HTA bodies where considerations other than QALYs are assessed and weighted qualitatively. In most deliberative processes, the relative weights given to the elements of value may remain implicit.

The principal approaches are considered in more detail in Table 2, which highlights some key issues and advantages of each, as well as common challenges.

Table 1: Approaches to the aggregation of overall value; issues and merits of each; and implications for the identification of the value-based price

	How is value aggregated?	Key issues specific to this approach	Key merits of this approach	Issues common to all approaches
Net benefit	As the sum of the benefits, each assessed in monetary terms	Challenges estimating the value in monetary terms of each type of value Allocating a monetary value to health has been always one of the mayor criticisms	Arguably, a better grounding in economic theory. Facilitates the comparison of value and value for money across health and other sectors Use of monetary value may resonate better with some (private) payers	A consensus on the perspective (NHS? government? societal?) from which value is assessed is required, regardless of which approach is used. The metrics by which aspects of value other than health are *measured* needs to be defined, as a prior step to valuing them
MCDA	As the sum of the points assigned to each aspect of value	The cost-effectiveness threshold would need to be reassessed in terms of the cost per incremental "point"	A pragmatic approach, widely used in the UK public sector. A more transparent (than a weighted QALY or deliberative process alone) means of addressing multiple criteria. MCDA is used in local NHS commissioning; potential to develop a consistent priority setting framework for both new and existing health care technologies	
Weighted adjusted QALYs	1. By QALYs gained, up-rated or down-rated by one or multiple weights to represent the magnitudes of other aspects of value; or	Assumes that all other sources of value are proportional to the number of QALYs gained. There are implications for the threshold. If the value of new technologies is assessed in terms of a range of criteria, then opportunity cost also has to be considered in the same terms, not	Is it relevant to state here the classic arguments in favour of the QALY such as: - Allows for comparisons across therapeutic areas in the NHS - "A QALY is a QALY" argument - Well established in the UK within HTA bodies (and academic centres)	

	How is value aggregated?	Key issues specific to this approach	Key merits of this approach	Issues common to all approaches
	2. Direct estimation of how people trade off QALY gains with other value elements	just QALYs foregone. Even if a simple social weighting or QALYs is applied, the opportunity cost will change.	- Understood by health economics community	
Deliberative process	Weights are assigned by a committee to each relevant aspect of value	Weights are often implicit Implications for the threshold	Provides an element of flexibility Well-recognised approach used by HTA bodies around the world	

Source: Adaptation of Sussex, Towse and Devlin (2012)

4. Attribution of Value in a Diagnostic-Treatment Combination

When a treatment and a diagnostic are used in combination to target a subgroup of patients, the value created is a "joint product" as defined in economics; the total value created depends on the combination and the attribution of some portion of the value to one or the other is essentially arbitrary. This presents a challenge to the concept and operationalization of VBP for both diagnostics and treatments.

We can illustrate this arbitrariness by considering an extreme situation in which (1) the treatment cannot be used without the test (let us assume the adverse effects are very high for the "wrong" patients) and (2) the test has no other application. Together a value of, say, 100, is created. If the test is taken away, the treatment has zero value. If the treatment is taken away the test has zero value. In other (less extreme) situations, the test may increase the value of the drug by enabling it to be targeted: let us assume, for example, that the net benefit to the health system of the drug on its own is 60 and with the test the net benefit increases to 100. Therefore the test adds a value of 40. The test also has some value in the absence of the drug as it can be used to help target treatments that are much less effective. On the other hand, suppose the value of the test without the drug is 20. The drug increases the value to 100 and so adds 80 to the value of the test on its own. Thus we can see that there is no "correct" way of dividing the joint value (of 100) of the test and drug between them. We can allocate the benefit using a rule, (and we have illustrated two), but it is essentially arbitrary.

This paper cannot "solve" this thorny theoretical problem, but some key elements to consider can be identified. Garrison and Austin (2007) have pointed out that how value is allocated across patients, payers, diagnostic manufactures and drug manufacturers (the "value capture") depends on the institutional context -- for example, whether the drug treatment was priced before the diagnostic was available, the relative strength of intellectual property protection for drugs and diagnostics, whether pricing and reimbursement of the medicine or diagnostic is flexible or administered, and other factors. We have shown in our simple example above that one rule is to look at which comes first and then allocate to the other one the residual of the joint value (recognising that the value attribution is different depending on which one comes first).

Who "captures" this value influences how much R&D is undertaken and therefore whether value is likely to be created in the first place. In this context, VBP principles should pay close attention to "dynamic" as well as "static" efficiency. Static efficiency is concerned with whether a treatment or diagnostic is cost-effective given current prices and usage patterns. Dynamic efficiency is concerned with how P&R policies influence the incentive to innovative: do they encourage the optimal rate of innovation?

In the case of an "at-launch" diagnostic-treatment (Dx-Tx) combination then, providing that overall value is identified, assessed and rewarded by payers, both drug and diagnostic manufacturers have the potential to make appropriate commercial arrangements with each other to maximise their joint opportunity for creating value[1]. However, this may be more difficult when a diagnostic test alone is being considered. This may be the case when a new test may be able to increase the overall joint value of an existing Dx-Tx combination because of, say, greater accuracy yielding fewer false positives and false negatives.

When it is possible to develop different platforms or versions of the same test (for example, hospital laboratories can create "in-house" toolkits for the same marker), there is a risk of "class effect" reimbursement recommendations (Drummond, Griffiths and Tarricone, 2009) whereby clinical and cost-effectiveness data of one test are extended to other versions of the test, without sufficient evidence. The key issue is that follow-on versions may have different technical characteristics compared to the first-in-class test, including accuracy, which has a significant impact on the overall health gains of Dx-Tx combinations. This makes generalisation of evidence across tests flawed. There is a need to verify and fully recognise the variability among tests with the same clinical use and the implications for the incremental value delivered by each of them.

If the rewards for targeting such a diagnostic are not sufficient to support appropriate evidence development, then the development and use of the test may be suboptimal. As Garrison and Austin (2007) noted, to the extent that intellectual property rights (IPR) are weak, and it is relatively easy and inexpensive to develop "follow-on" tests, then the market forces will be similar to those for a class of generic drugs. Price will be driven down to marginal production and distribution cost. This

[1] This may not happen. For a discussion of some of the issues using a theoretical model of a relationship between a "research unit" (the diagnostics company) and a "customer" (the pharmaceutical company), see Aghion and Tirole (1994).

will not provide sufficient incentive for "first-in-class" tests to produce the optimal amount of supporting evidence. VBP may need to be supplemented with other incentives, such as data or marketing exclusivity or public subsidies (for example, to fund evidence generation), to encourage socially optimal levels of innovation in diagnostic testing. Further discussion around IPR issues for diagnostics is beyond the remit of this paper.

5. Three Examples of Institutional Processes for Diagnostics

Historically, pricing and reimbursement systems for diagnostics have focused on costs (Garrison and Austin, 2007). This has meant that the price of a new diagnostic is fixed based on the price of existing tests with similar clinical use or similar characteristics, or based on production cost. For example, in the US, a number of diagnostics are reimbursed through a combination of reimbursement codes describing laboratory protocol stages (Gustavsen, Phillips and Pothier, 2010). There is an emerging tendency among countries, such as the UK and Australia, however, to extend HTA arrangements to diagnostic tests.

5.1 The UK system

In the UK, NICE was established to provide an independent assessment of the cost-effectiveness of medical technologies to guide decision making in the NHS. In 2009-2010 the Diagnostics Assessment Programme (DAP) was created to assess diagnostic technologies within NICE's remit (NICE, 2011).

DAP's responsibility includes genetic tests with a medical purpose. It focuses mainly on "stand-alone" diagnostics. Companion diagnostics, which identify subpopulations that respond best to a new drug, usually are assessed alongside the pharmaceutical within a NICE Technology Appraisal.

DAP recognises that the evaluation of diagnostics differs from that of treatments, mainly because diagnostics do not have a direct impact on health outcomes. However, the current DAP approach does not allow the decision maker to consider a broad set of outcomes, including the value of information on patients' conditions independent of health gains. This is because the current method very closely follows that used for medicines; the measure of patient benefit is based purely on the QALY.

In the case of companion diagnostics assessed in conjunction with treatments, the incremental value offered by each of the two technologies is an issue, as discussed in section 4. In the NICE appraisal of trastuzumab for the treatment of early-stage HER2-positive breast cancer (NICE, 2006), it was accepted that an assay had to be used to identify the relevant patient subpopulation, according to the marketing authorisation. The guidance states that the cost of HER2 testing was included in the economic analysis, but it did not explicitly include the amount as it did for the treatment. Furthermore, there was no explicit mention of test specificity and sensitivity.

5.2 The Australian system

Australia currently has a dedicated HTA process for new diagnostics including both stand-alone and companion diagnostics. Diagnostics are classified as "medical services". The Medical Services Advisory Committee (MSAC) advises the Ministry of Health as to listing on the Medicare Benefits Schedule (MBS), which is separate from the Pharmaceutical Benefits Scheme (PBS). Hence, currently, companion diagnostics and associated treatments are assessed through different committees (MSAC and PBAC, respectively) with no clear structure for consideration of the interactions between or benefits from joint use.

From November 2012, a new coordinated process will be implemented for "co-dependent technologies".[1] An "integrated" application combines information developed by the diagnostic manufacturer, the drug manufacturer, and by both. Different funding programs mean, however, that listing decisions for diagnostics and treatments included in the "co-dependent technologies" category still will be made separately by PBAC and MSAC.

An important issue relates to the draft December 2010 guidelines on evidence requirements for co-dependent technologies. It proposes a new evidence hierarchy to demonstrate the clinical benefits of tests which includes, as the preferred option, a patient randomisation to use of test ("direct

[1] For more details see DHA (2012).

evidence"). However, not considered is that the choice of the most adequate study type and design should be informed by an explicit consideration of the added value generated by each type of evidence compared with the cost and feasibility of collection.

5.3 The US system

The arrangements in US managed care organizations to assess drugs and diagnostics vary. In most cases, payers have formulary committees that consider the value of new drugs, but no similar arrangements for assessing the value of diagnostics. Diagnostic reimbursement varies widely (Housman, 2011). Diagnostics are generally reimbursed in both the public and private sectors on a crude cost-based coding system. Government payers -- Medicare and Medicaid -- have reimbursement levels that vary by state. Test manufacturers must negotiate with private payers individually by test and procedure code. The reimbursement by code within a private payer typically ranges from 60% to 110% of the Medicare reimbursement by procedure (Gustavsen, Phillips and Pothier, 2010). Specific tests usually do not have a unique code and usually cost less than $500.

If the standard coding system can be avoided, the US health care system may be willing to pay for at least some elements of value. Oncotype DX was not developed as a companion diagnostic, being launched independently of a chemotherapy treatment, thereby putting the requirement to secure value based reimbursement squarely on the assay manufacturer.

Under the current procedure-based coding mechanism, the 21-gene assay, using the procedure-based code stacking approach, would have totalled approximately $580 using a Medicare fee schedule basis (Gustavsen, Phillips and Pothier, 2010).

Instead, the manufacturer pursued a value-based pricing model utilising diagnostic clinical trial and patient outcome studies to demonstrate clinical differentiation and cost-effectiveness when the 21-gene assay was utilized for node-negative breast cancer patients. The main focus of cost-effectiveness argument was the cost-offset obtained by not undertaking expensive chemotherapy treatment for women at low risk of disease recurrence. As the first in the market to utilize this model, they were able to achieve reimbursement for the assay at roughly seven times the code-stacking reimbursement (i.e. around $3,500). This was by no means a trivial or quick undertaking: it took over four years to obtain nearly 90% payer coverage (Gustavsen et al., 2010).

Value based pricing for this test was aided by its first-mover advantage, the importance of the cost-offset (as opposed to the value of any health effects), and investment in data collection and publication.

5.4 Proposed institutional processes for diagnostics

There are several factors to trade-off in designing institutional arrangements.

First, it is important to build up experience of dealing with HTA for diagnostics as there are learning effects linked to cumulative experience. The particular issues for diagnostics include:

1. Generating and interpreting evidence on aspects of benefit, such as information for patients, that are less likely to occur in drug appraisal
2. Recognising the different circumstances for feasible study design and evidence collection for diagnostics as compared to drugs (for a discussion see Drummond, Griffiths and Tarricone, 2009; Taylor and Iglesias, 2009). If pricing and reimbursement systems do not capture the full benefits brought to society by diagnostics and there is no sufficient protection of IPR, diagnostic manufacturers will not have incentives to invest in evidence development to raise the standard of clinical data available to support the case for using a test
3. The specific incremental characteristics of competitive tests with similar clinical use and how these may or may not translate into incremental value or cost savings for the payer

Second, there is a need to consider the possible economies of scale of having a separate committee for diagnostics, which could increase throughput. At the same time, it is critical to achieve synergies across drugs and diagnostics in three respects:

1. there are economies of scope from one group dealing with both drugs and diagnostics
2. the joint product nature of an “at launch” combination requires one group to review both technologies in one package
3. the health system should be looking for the same value across all technologies, which requires a consistent approach to willingness to pay for value

This suggests that there is a case for two types of institutional arrangement: (1) a separate diagnostics committee to develop and use diagnostics-specific expertise -- however, there may be a

trade-off if there are not enough decisions to justify a distinct committee and (2) a joint drug-diagnostic review of “at launch” technologies, logically done by the drug committee, to exploit synergies across diagnostics and drugs. However, most drug committees lack of expertise in the diagnostics area. This could be addressed by involving a sub-group of the diagnostics committee in any deliberations of the drug committee on drug-test combinations or by overlapping membership.

Which route is followed for the appraisal of a diagnostic test would depend on the following characteristics:

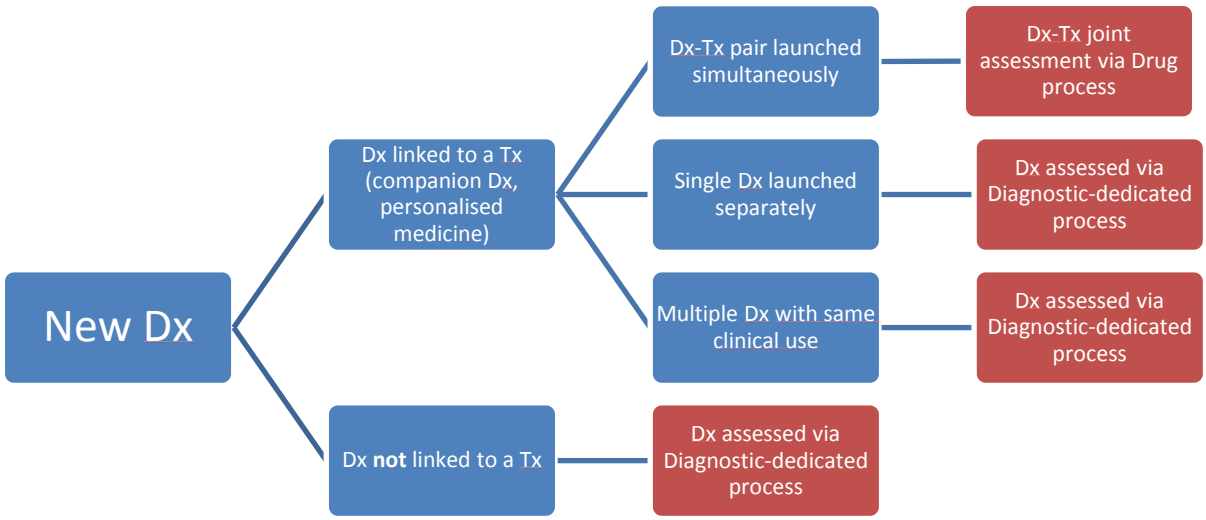
1. *Purpose -- in particular the key distinction is between companion diagnostics, and other diagnostics not directly linked to treatments.* The latter includes those used for screening purposes and those used to ascertain future clinical events and outcomes (prognostic information)

2. *Timing of launch with respect to the corresponding treatment*, in the case of companion diagnostics; in particular, whether a new diagnostic is developed and launched alongside the treatment or whether it enters the market separately (before or after the treatment)

3. *Presence of competitive tests with similar purposes, but with a different cost and/or quality profile*. If there are alternative ways of delivering the companion diagnostic, their assessment should be separate from the one used for the joint assessment. A comparative analysis between existing tests with similar clinical use would lead to optimal decisions.

Figure 2, below, illustrates how those characteristics could drive the selection of the process for a new diagnostics.

Figure 2. Institutional processes for the assessment of value of new diagnostics



Other situations related to the evolution of the market for treatments linked to diagnostics can arise. For example, second-in-class medicines using an existing test may be developed such as lapatinib, which employs the same testing regimen as trastuzumab for selecting women with metastatic HER2-positive breast cancer. In those cases, the assessment of two or more drug-test combinations could be done via the drug process review with a focus on the comparison between the treatments outcomes.

It would be essential that both diagnostic-dedicated and drug processes use a common, comprehensive approach to assessing value, using the one of the approaches to weighting value that is set out in Table 1 above. For example, if an MCDA approach is used to assess the value of drugs, it also should be used to assess the value of diagnostics, ideally using a consistent set of weights to recognise value from whatever source.

6. Conclusions

We have argued for a value based approach for pricing and reimbursement that reward innovation in diagnostics and drugs. We have set out possible institutional arrangements to support this. These include sending combined “at-launch” drug-test combinations to a drug assessment committee and establishing a separate specialist committee to review diagnostic tests that lie outside of an at-launch situation. NICE is heading in this direction, but does not yet have a comprehensive approach

to assessing the value of diagnostics or drugs. In Australia, the common methodology needs to be supported by synergies in decision making and a realistic view of evidence development. In the US, an important precedent has been set by Oncotype DX for pricing for a diagnostic by value (as opposed to by cost). However, evidence was primarily around cost offset rather than health gain, and it has taken several years for the test to achieve comprehensive cover. Both public and private sector payers need to bring a value-based approach and specialist expertise to diagnostic reimbursement decisions.

A value-based approach to pricing is necessary, but not sufficient, to stimulate the development of new diagnostic tests. Issues such as IPR also may have to be addressed if sufficient evidence is to be generated to meet the requirements of payers.

References

Aghion, P. and Tirole, J. (1994) The Management of Innovation. *The Quarterly Journal of Economics.* 109(4), 1185-1209.

Ash, D., Patton, J. and Hershey, J. (1990) Knowing for the sake of knowing: The value of prognostic information. *Medical Decision Making.* 10(1), 47-57.

Claxton, K. and Posnett, J. (1996) An economic approach to clinical trial design and research priority-setting. *Health Economics*. 5(6), 513-524.

Danzon, P. and Towse, A. (2002) The economics of gene therapy and of pharmacogenetics. *Value in Health.* 5(1), 5-13.

Devlin, N. and Sussex, J. (2011) *Incorporating multiple criteria in HTA: Methods and processes*. London: Office of Health Economics.

DH (Department of Health) (UK). (2010) A new value-based approach to the pricing of branded medicines -- a consultation. London: Department of Health. Available at http://www.dh.gov.uk/en/Consultations/Liveconsultations/DH_122760. [Accessed 1 April 2012].

DHA (Department of Health and Ageing. Australia). (2012) *Co-dependent and hybrid technologies.* Health Technology Assessment Access Point. Available at: http://www.health.gov.au/internet/hta/publishing.nsf/Content/co-1 [Accessed 25 April 2012].

Drummond, M., Griffiths, A. and Tarricone, R. (2009) Economic evaluation for devices and drugs -- Same or different? *Value in Health.* 12(4), 402-406.

Ellsberg, D. (1961) Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics.* 75(4), 643-669.

Garau, M., Shah, K.K., Mason, A., Wang, Q., Towse, A. and Drummond, M. (2011) Using QALYs in cancer: A review of the methodological limitations. *Pharmacoeconomics.* 29(8), 673-685.

Garrison, L.P. and Austin, M.J.F. (2007) The economics of personalized medicine: A model of incentives for value creation and capture. *Drug Information Journal.* 41(4), 501-509.

Gustavsen, G., Phillips, K. and Pothier, K. (2010) *The reimbursement landscape for novel diagnostics.* Weston, MA: Health Advances.

Han, P., Klein, M.P. and Arora, N.K. (2011) Varieties of uncertainty in health care: A conceptual taxonomy. *Medical Decision Making.* 31(6), 828-838.

Housman, L. (2011) Innovative tests for patient-tailored treatments: Demonstrating value in the age of personalized medicine. *Personalized Medicine.* 8(2), 183-189.

Kenen, R. (1996) The at-risk health status and technology: A diagnostic invitation and the 'gift' of knowing. *Social Science & Medicine* 42(11), 1545-1553.

Lee, D., Neumann, P. and Rizzo, J. (2010) Understanding the medical and nonmedical value of diagnostics testing. *Value in Health.* 13(2), 310-314.

Neumann, P.J., Cohen, J.T., Hammitt, J.K., Concannon, T.W., Auerbach, H.R., Fang, C. and Kent, D.M. (2012) Willingness-to-pay for predictive tests with no immediate treatment implications: A survey of US residents. *Health Economics.* 21(3), 238-251.

NICE (National Institute for Health and Clinical Excellence). (2011) *Diagnostics assessment programme manual*. Manchester: NICE.

NICE (National Institute for Health and Clinical Excellence). (2006) Trastuzumab for the adjuvant treatment of early-stage HER2-positive breast cancer. Technology Appraisal 107. London: NICE.

Shah, K.K. (2009) Severity of illness and priority setting in healthcare: A review of the literature. *Health Policy*. 93(2-3), 77-84.

Sussex, J., Towse, A. and Devlin, N. (2012) Operationalising value based pricing of medicines: A taxonomy of approaches. *Pharmacoeconomics.* (forthcoming).

Taylor, R. and Iglesias, C. (2009) Assessing the clinical and cost-effectiveness of medical devices and drugs: Are they that different? *Value in Health.* 12(4), 404-406.

Walker, I. and Newell, H. (2009) Do molecularly targeted agents in oncology have reduced attrition rates? *Nature Reviews Drug Discovery.* 8, 15-16.

Young, D. (2011) Experts: Pfizer-Abbott NSCLC Rx/Dx delivers on personalised medicine promise. *Scrip Intelligence.* 31 Augus5t.