
Book Section:
Drummond, M.F. orcid.org/0000-0002-6126-0944, Tarricone, Rosanna and Torbica, A. (2018) Economic Evaluation of Medical Devices. In: Hamilton, J.H., (ed.) Oxford Research Encyclopedia of Economics and Finance. Oxford University Press , Oxford

https://doi.org/10.1093/acrefore/9780190625979.013.105

ECONOMIC EVALUATION OF MEDICAL DEVICES*

Michael Drummond[1]

Rosanna Tarricone[2,3]

Aleksandra Torbica[2,3]

1. Centre for Health Economics, University of York, UK.
2. Department of Policy Analysis and Public Management, Università Bocconi, Milan, Italy.
3. Cergas, SDA Bocconi School of Management, Milan, Italy.

* Article prepared for *Oxford Encyclopaedia of Economics and Finance*

Abstract

There are a number of challenges in the economic evaluation of medical devices (MDs). They are typically less regulated than pharmaceuticals and the clinical evidence requirements for market authorization is generally lower. There are also specific characteristics of MDs, such as the device-user interaction (learning curve), the incremental nature of innovation, the dynamic nature of pricing and the broader organizational impact. Therefore, a number of initiatives need to be taken in order to facilitate the economic evaluation of MDs. First, the regulatory processes for MDs need to be strengthened and more closely aligned to the needs of economic evaluation. Secondly, the methods of economic evaluation need to be enhanced, by improving the analysis of the available clinical data, by establishing high-quality clinical registries and better recognizing MDs' specific characteristics. Thirdly, the market entry and diffusion of MDs needs to be better managed, by understanding the key influences on MD diffusion and by linking diffusion with cost-effectiveness evidence through the use of performance-based risk-sharing arrangements.

1. Introduction

Medical devices (MDs) represent an important group of health technologies. In 2011 it was estimated that there were more than 200,000 MDs on the European market (Fraser et al, 2011). However, MDs represent a very heterogeneous family of technologies, consisting of both diagnostic and therapeutic devices, and ranging from 'low risk' devices such as scalpels, to 'high risk' devices such as cardiac implantable electrical devices. While some devices only require very simplified assessment, others need to be assessed through a full evaluation of safety, efficacy, effectiveness and cost-effectiveness.

Many of the early applications of economic evaluation in health care were of the use of medical procedures involving devices, such as kidney dialysis (Klarman et al 1968), or of expensive diagnostic devices such computed tomography (CT) scanners (Office of Technology Assessment, 1978). However, the more recent literature in economic evaluation has focused largely on pharmaceuticals, reflecting the growth of formal requirements for new drugs to demonstrate cost-effectiveness in order to be reimbursed in many jurisdictions (Drummond, 2013). For example, in 2016 the National Institute of Health and Care Excellence (NICE) in the UK, one of the most active heath technology assessment organizations worldwide, published 50 technology appraisals of pharmaceuticals, but only 3 appraisals of medical devices and 6 of diagnostic technologies (www.nice.org.uk).

This relative lack of economic evaluations of MDs as compared with pharmaceuticals has led some authors to question whether devices have any special characteristics that inhibit their assessment and whether, over time, the methods guidelines for economic evaluation have become more suited to the assessment of pharmaceuticals (Tarricone et al, 2017a). The challenges in conducting economic evaluation of medical devices are related to two major issues. The first lies in the policy domain and regards current regulatory requirements, which are typically less stringent than for pharmaceuticals. As a consequence, evidence produced by manufacturers applying for marketing approval is often not sufficient for subsequent economic evaluation analysis. The second issue is purely methodological, and reflects the ongoing debate on whether devices have distinctive features that require different or modified methods for economic evaluation to be applied properly. This article discusses the challenges in conducting economic evaluations of MDs, from both policy and methodological viewpoint, and provides some suggestions on how they might be addressed.

1.1 Regulatory requirements for MDs

One critical issue relates to the fact the regulatory requirements for devices are typically less stringent than those for pharmaceuticals. In order to obtain approval to market, pharmaceuticals normally require two adequately controlled clinical trials. In contrast, the level of clinical evidence required for the marketing of MDs varies by the risk classification of the device. The risk classifications differ slightly by jurisdiction, but generally divide devices into 3 or 4 groups (Tarricone et al, 2014).

Low risk devices, such as crutches and scalpels, are generally exempt from premarket approval, although they need to be registered and are subject to labeling regulations. In the European Union this means obtaining a CE (Conformité Europeéne) mark. Devices of medium risk, such as catheters and infusion pumps, do require premarket approval and often the manufacturer is required to submit a

dossier of relevant clinical and non-clinical data. The data required usually depends on the level of perceived risk. For example, in some jurisdictions (eg the European Union), orthopaedic implants are classified as 'medium risk' but manufacturers would be expected to provide some clinical data and also establish a registry so that the performance of the device could be monitored after it enters the market.

Finally, high risk devices such as coronary stents, cardiac pacemakers and implantable defibrillators always require clinical data, although the nature of that data may differ by jurisdiction. In the US the Food and Drug Administration often requires a randomized controlled trial, similar to its requirement for pharmaceuticals, whereas in the EU it may be possible to obtain market approval by conducting an observational study (eg a clinical registry). This was the case for transcatheter aortic valve implantation (TAVI), where an RCT was conducted in the US, but an observational study was established in Europe (Tarricone and Drummond, 2011).

Another feature of the regulation of MDs that differs from pharmaceuticals is that manufacturers can claim that their device is 'substantially equivalent' to another marketed device for which clinical data are available. An example of this approach is the 510(k) process in the US (Sorenson and Drummond, 2014). This can greatly assist some manufacturers (known as 'fast followers') but has the effect of not allowing 'data exclusivity' to the manufacturer of the original device, as is the case for pharmaceuticals. This can have the effect of reducing the incentives for manufacturers to invest in clinical research, if that research is not mandatory, since the data generated can also be used by their competitors.

A major impact of the current processes for regulating devices is that there is a relative lack of clinical data at the time of product launch, although there remains the possibility of further clinical data collection while the device is in regular clinical use. This limits the possibilities for conducting economic evaluations early in the life cycle of the technology, especially at the time of product launch. Therefore, it has been argued that a different strategy may have to be followed for conducting economic evaluations of MDs, focusing more on assessments in the post-marketing phase (Tarricone et al, 2014). However, conducting economic evaluations in real life settings poses methodological challenges in the analysis of observational data, which are subject to numerous biases. Furthermore, evidence synthesis becomes even more demanding when data needs to be synthesised across studies with different designs (Schnell-Inderst et al, 2017).

1.2 Special characteristics of MDs

It has been suggested that there are a number of special characteristics of MDs that pose additional challenges for their clinical and economic evaluation (Drummond et al, 2009). These include the fact that (i) there is often an interaction between the device and the user (eg a surgeon), meaning that the performance of the device may change over time (the so-called 'learning curve'), as a given user gains more experience (ii) MDs are characterized by incremental innovation, where small improvements are made, either through modifications to an existing device or the introduction of new products (iii) the introduction of MDs can have a substantial organizational impact, by requiring training or investments in infrastructure and (iv the pricing of MDs is more dynamic than that for other health technologies, owing to the extensive use of competitive tendering in the procurement process. In particular, the learning curve may have substantial effects on both effectiveness and costs. A more experienced surgeon may obtain better clinical results than one performing a procedure for the first time. In turn, better clinical outcomes might reduce the total

length of hospital stay, or the need for subsequent procedures, which could also have an impact on costs. Also, a more skilled surgeon may perform the initial procedure more quickly, which could reduce the amount of expensive operating theatre time required.

There is some evidence to support these assertions. For example, the learning curve for MDs has been observed in several clinical trials. In the CLASICC Trial of laparoscopic-assisted colorectal surgery, the proportion of patients requiring 'revisions' (ie repeat procedures) fell from 38% in the first year of the trial, to 16% in year six (Guillou et al, 2005). Similarly, in a clinical trial of spinal chord stimulation versus percutaneous myocardial laser revascularization for patients with angina, improvements in the outcome from use of the new procedure caused the estimated incremental cost-effectiveness ratio (ICER) to fall from £230,000 per quality-adjusted life-year (QALY) in 2000/1 to £18,000 per QALY in 2002/3 (Dyer et al, 2008). Learning curve effects have also been observed in the use of procedures in regular clinical practice. Using individual patient level data obtained from a hospital discharges database in Germany over the period 2006-13, Varabyova et al (2017) found that, over time, both the in-hospital mortality and length of stay became lower for patients undergoing endovascular aneurism repair (EVAR), implying that the impact of the learning curve on cost-effectiveness can be more substantial than that on effectiveness alone. Conversely, there was no similar learning effect for a related procedure, fenestrated endovascular repair (fEVAR), possibly because much of the 'learning' from performing EVAR procedures was transferable to fEVAR.

The issue of incremental innovation was raised in a clinical and economic evaluation of stapled hemorrhoidectomy conducted for the National Institute for Health and Care Excellence (NICE) in the UK (Burch et al, 2008). The researchers found that, although clinical trials comparing staple guns versus conventional surgery did exist, most of the trials were conducted using older versions of the gun, which had undergone several modifications. However, it could be argued that, as long as the modifications represented improvements to the device's performance and the learning curve in using the new version of the gun was not substantial, it would be conservative to use the older clinical evidence in the economic evaluation. Nevertheless, in some cases using clinical data from an older version of a technology could have a substantial impact on the results of the economic evaluation.

Organizational impacts have been observed, but not examined quantitatively. For example, Tarricone and Drummond (2011) discussed the fact that adoption of transcatheter aortic valve implantation (TAVI) in a given hospital would normally involve investments in infrastructure and training. Because of the complexity and high risk of patients, TAVI would ideally need to be performed in a hybrid setting with equipment and facilities including both operating rooms and catheterisation laboratories (cathlabs). Alternatively, it could be delivered either in operating rooms with adequate visualisation equipment or in sterile cathlabs. This remains a challenge today, since 'pure' hybrid operating rooms or appropriately equipped operating rooms and cathlabs are the norm only in the most advanced hospitals and would need to be developed further if TAVI were to be more widely diffused. Logistics and multi-disciplinary team approaches also represent an organizational issue for many hospitals and must be assessed against investment costs in training and equipment. Moreover, a minimum number of cases needs to be planned in order to secure the return on the investment and, more importantly, the sufficient level of performance.

The impact of price changes on the cost-effectiveness of MDs was illustrated in the technology appraisals conducted on drug-eluting stents (DES) by the National

Institute for Health and Care Excellence (NICE) in the United Kingdom (UK). In the initial appraisal, conducted in 2004, NICE found that DES were cost-effective (according to its threshold of £20,000 per QALY gained) for patients with a high risk of restenosis. However, when NICE reappraised DES in 2008, the price of bare metal stents (the alternative technology) had fallen substantially, pushing the estimated incremental cost-effectiveness ratio (ICER) above NICE's threshold (Drummond et al, 2009).

Therefore, taken together, these special characteristics can influence the planning and analysis of cost-effectiveness studies. For example, the potential existence of a learning curve suggests that the cost-effectiveness of MDs should be studied over a period of time, as the results when the device is first introduced may be misleading. The potential existence of incremental innovation is a challenge for the clinical component of the assessment, as (in principle) separate clinical studies may need to be conducted for each version of the device. In addition, if the changes in the benefits to patients from using the new version of a device are small, these may not be reflected in changes in the generic instruments used for estimating the quality-adjusted life-years (QALYs) gained. Therefore, analysts wanting to estimate the value of these changes may need to estimate individuals' willingness-to-pay or conduct a discrete choice experiment (Wilkinson and Drummond, 2015). (See Section 2.2.3 below.)

The potential for organizational impacts of devices suggests that the perspective for costing needs to be broad, covering a wide range of hospital departments and possibly also primary care, if use of the device changes the pattern of care. In addition, implementation costs, such as training and the provision of additional support services, may also need to be considered. Finally, the possibility of dynamic pricing suggests that a sensitivity analysis should be performed on both the acquisition cost of the new device and the one that it is replacing.

A more fundamental point relating to the special characteristics of MDs is that, as compared with drugs, the 'real world' cost-effectiveness may differ substantially from that observed within the context of controlled clinical trials. This suggests that the conduct of economic evaluations might be better focused on the post-marketing phase, rather than pre-marketing phase, as in the case of pharmaceuticals. This issue is discussed further in Section 3 below.

On the other hand, it has been argued that, to the extent that these differences exist between MDs and other technologies (in particular pharmaceuticals), the differences are only a matter of degree (Taylor and Iglesias, 2009). For example, over time physicians may learn how to prescribe a drug more precisely in order to obtain the optimum balance between efficacy and adverse events, new modes of administration for a given drug (eg oral instead of intravenous infusion) might be viewed as incremental innovations, or alternatively may lead to shifts in care from an inpatient to an outpatient setting with the associated changes in infrastructure. Furthermore, tendering or price competition applies to pharmaceuticals in some jurisdictions (eg US managed care) and there are tendering processes for vaccines in many jurisdictions. Nevertheless, the special characteristics mentioned above are likely to have a greater impact in the economic evaluation of devices than other health technologies, so they need to be considered.

Another important characteristic is that many MDs, such as CT scanners, are diagnostic. The economic evaluation of these technologies presents three specific challenges. First, the major benefit arising from diagnostic technologies is their ability to improve treatment choices and hence final health outcomes. Therefore, an

accurate diagnostic test is only as good as the treatment that follows it. This means that the economic evaluation of a diagnostic technology usually encompasses the evaluation of the associated treatment technologies, with all the inherent challenges of conducting the required clinical evaluations.

Secondly, although the major economic benefit of diagnostic technologies results from the improvement in treatment outcomes, there may be value in just knowing the accurate diagnosis, both to the patient and the physician, even if this does not affect the immediate treatment choices. In some cases there may be value from the reassurance offered by a diagnostic test. Although these aspects of value could, in principle, be detected by the standard measures of outcome used in economic evaluations, such as the quality-adjusted life-year (QALY), on occasions these measures may lack the adequate sensitivity. Therefore, it might be necessary to use alternative measures of value, such as willingness-to-pay (Lin et al, 2013).

Thirdly, there are indivisibilities in the use of large diagnostic MDs such as scanners. This means that, not only is it necessary to show that individual indications for use of the scanner, it is also necessary that the totality of the workload required to occupy the scanner results in a cost-effective use of resources. If there were only a few potentially cost-effective indications for the scanner, this would not justify the large capital expenditure.

2. Initiatives to improve the economic evaluation of MDs

2.1 Improving the clinical evidence base for MDs

It was mentioned above that the paucity of clinical data, particularly prior to market entry of MDs, can limit the possibilities for economic evaluation and health technology assessment (HTA). Drummond et al (2016) discussed several options for resolving this problem. First, the requirements for conducting clinical studies in the pre-marketing phase could be strengthened by, for example, insisting on randomized controlled trials for all high-risk devices and some medium risk devices (eg orthopaedic implants). However, this would impose greater costs on manufacturers and delay access to new treatments (Shuren and Califf, 2016). However, these considerations need to be balanced by the fact that potentially unsafe or ineffective devices could be allowed on the market before adequate evidence is available (Food and Drug Administration, 2016).

Another initiative would be to tighten the conditions under which a manufacturer can claim ‘substantial equivalence’ to a pre-existing device already on the market. This would go some way towards restoring the incentives for research that follow from maintaining data exclusivity. However, this could result in a waste of resources in conducting clinical studies that are not necessary. Also, since some of the ‘fast-followers’ are small and medium-size enterprises, which might find the cost of conducting clinical studies a challenge, this policy would have the impact of reducing price competition, thereby denying the health care system of some efficiency gains.

The third possibility would be to strengthen the arrangements for post-market research. One argument for this policy would be that, due to some of the special characteristics of MDs mentioned above, requiring more RCTs in the pre-market phase might not be very helpful. For example, if there is likely to be a substantial learning curve for the device in question, it would be preferable to base the economic evaluation on clinical effectiveness in regular use, rather than under the experimental

conditions of an RCT. Also, it there are likely to be several modifications of the product in the early years of its use, the clinical data gathered in the trial may quickly become outdated.

However, there are several disadvantages of delaying economic evaluation to the post-market phase. It is often more difficult, and costly, to remove a treatment or technology from the market once it has entered the health care system. Also, once the technology is routinely available it is more difficult, if not impossible, to conduct a randomized clinical trial, since this implies that 50% of patients participating in the trial would not receive the new treatment. However, ways might be found to overcome this by, for example, randomizing one group of patients to delayed treatment rather than no treatment.

These difficulties mean that the clinical studies conducted post-launch are likely to be observational studies, which often suffer from selection bias and are therefore more likely to produce a biased assessment of relative treatment effect (Grieve et al, 2016). Of course statistical approaches are available to minimize the impact of selection bias, but it cannot be eliminated. However, in some situations only observational studies may be possible.

These considerations suggest that any strategy for conducting economic evaluation in the post-marketing phase needs to be part of a broader strategy for the managed entry of MDs, involving decisions about (i) which MDs should be prioritized for further research (ii) the design of the clinical and economic evaluations and (ii) the arrangements for pricing and reimbursement of the device, both while under further study and after the results of the subsequent clinical and economic evaluation become available. (This is discussed further in Section 3 below.)

2.2 Improving economic evaluation methods

Two recent systematic reviews shed some light on the current status of the economic evaluation of devices. Ciani et al (2017) analysed a sample of HTA reports in the cardiovascular field in order to assess whether there were any key differences in the methods employed in studies of MDs, as compared with studies of pharmaceuticals. They found that there were several differences: (i) in the types of clinical studies forming the basis of the HTAs (with a greater reliance on observational studies, and less reliance on RCTs in the case of MDs); (ii) how the health problem and use of the technology were considered; (iii) the description and technical characteristics of the technology and; (iv) the consideration of the organizational aspects of use of the technology. Most of these differences arose because of the relative complexities in the use of devices, in terms of the number of interacting components, such as the number of groups and organizational levels targeted by the intervention, the number and variability of the outcomes and the degree of flexibility or tailoring of the intervention.

In the second review, Tarricone et al (2017b) reviewed all the published economic evaluations and HTA reports for two cardiac devices, TAVI (an example of an emerging technology) and implantable cardioverter defibrillators (an example of a mature technology). The authors assessed each study using the CHEERS (Consolidated Health Economic Evaluation Reporting Standards) checklist (Husereau et al 2014), supplemented by some additional categories reflecting the special characteristics of MDs. In general, the methods used in the studies were fairly representative of the broader economic evaluation literature. However, the special characteristics of MDs were not very well addressed. A learning curve was considered in only 16% of studies on TAVI. Incremental innovation was more

frequently mentioned in the studies of ICDs, but its impact was considered in only 34% of the cases. Dynamic pricing was the most recognized feature, but was empirically tested in less than half of studies of TAVI and only 32% of studies on ICDs. Finally, organizational impact was considered in only one study of ICDs but in almost all studies on TAVI, but none of the studies of estimated the size of any impact.

2.2.1 Improving the analysis of the available clinical data

If the methods of economic evaluation of MDs are to be improved, progress needs to be made in compensating for the paucity of well-controlled clinical studies of devices and in addressing the special characteristics of MDs. There is much to do in both these areas, but several examples of good practice can be identified. First, in a recent study of total hip replacement, Schnell-Inderst et al (2017) used a method of (clinical) evidence synthesis that allows for the meta-analysis of RCT and observational data, using bias adjustment based on formal elicitation. They performed an elicitation exercise using methodological and clinical experts to determine the strength of beliefs about the magnitude of internal and external bias affecting estimates of treatment effect. They incorporated the bias-adjusted treatment effects into a generalized evidence synthesis, using both frequentist and Bayesian statistical models, calculating relative risks as summary effect estimates with 95% confidence/credibility intervals to capture uncertainty. They found that the pooled effect size strongly depended on the inclusion of observational data as well the use bias-adjusted estimates. The use of the bias-adjusted estimates shifted the pooled effect towards a lower treatment effect.

There still remains the challenge of producing unbiased estimates of treatment effect when only observational studies are available. In the economic evaluation more generally, either propensity scoring, a form of matching, and multivariate regression are frequently used. The methods either consider known confounders only, or also unknown confounders through the use of instrumental variables in the multivariate regression (Prentice JC et al 2014). Examples of the use of observational data to assess the cost-effectiveness of MDs include the study by Armeni et al (2016) on MitraClip for patients with moderate/severe mitral regurgitation and by Rognoni et al (2016) on transarterial radioembolization for patients with hepatocellular carcinoma. Both of these studies used the propensity scoring method.

A detailed discussion of the methodological considerations in using propensity scoring or multivariate regression to adjust for potential biases in observational studies is beyond the scope of this paper. However, it is worth noting that all these methods require data on the clinical characteristics of patients, for incorporation in the statistical model. These data may not always be available from administrative databases, or even some clinical registries. In addition, there may be unobserved factors that lead to bias in observational studies. In principle, the use of an instrumental variable can be useful, but the challenge is often to find a suitable instrument.

As was mentioned earlier, the most common clinical studies of MDs are clinical registries, conducted either in the pre-market or post-market phase. However, the vast majority of these registries study only a single device, or have very little information on factors that might be used as covariates in a multivariate analysis. Therefore is an urgent need is to improve the quality of registries for MDs, by making sure that there is treatment variation (eg including more than one device), classifying the risk categories of patients and, where possible, capturing data on resource use and the quality of life of patients (Tarricone et al, 2017c)

2.2.2. Estimating the learning curve

As was mentioned earlier, the learning curve for the use of some devices has been estimated, either within the context of a clinical trial, or by using an observational dataset, such as an administrative database. In the context of an economic evaluation of a new MD, an analyst can either use an estimate of the learning obtained for a similar device, or attempt to estimate the learning curve as part of the economic evaluation. An example of the first approach is the study by Varabyova et al (2017) cited earlier. The main issue in this case is whether previous estimates are good predictors of the likely learning curve for the new device. To this end it may be possible to group MDs in various categories, according to whether they are likely to have a shallow or a steep learning curve.

The learning curve has almost never been estimated as part of an economic evaluation. In the review by Tarricone et al (2017b), the learning curve was addressed in 16% of the studies on TAVI by varying the rates of complications and procedure success in a sensitivity analysis. Although there has been an extensive review of the impact of learning curves on outcomes (Ramsay et al, 2001), there have been very few assessments on the impact of learning curves on resource use or costs. See van Gestel et al (2013), Bell et al (2015) and Margier et al (2015) for recent examples.

2.2.3 Considering incremental innovation

Two issues are raised by incremental innovation. First, if the MD under study changes, this change may be substantial enough to render the previous clinical data out of date. This issue arises when product modifications are submitted to the regulatory bodies for approval. One analysis in the US showed that for 77 original market authorization applications for cardiac implantable electronic devices since 1979, the FDA has approved 5829 ‘supplements’ reflecting product modifications. In the vast majority of cases the FDA deemed that new clinical data were not necessary, although 37% of the supplements related to changes in the device’s design (Rome et al, 2014).

At the very minimum, analysts should note any product modifications that have occurred to the version of the device for which clinical data are available and comment on whether these are likely to impact on device performance. It may be concluded that many of the modifications are quite minor and likely only to impact of convenience of use, such as miniaturisation of the device. Others, such as an increase in battery life, may reduce costs if they reduce the need for replacing an implantable device. The impact of other modifications, such a change in the design of the device, could affect performance positively or negatively. Therefore data collection should be put in place to estimate any change, particularly if the design change is likely to mean that users face a new learning curve.

The other issue raised by incremental innovation is that the benefits to patients of a small change in the deign of an MD may not be detectable by generic preference-based quality of life measures, like the EQ-5D, that are used to estimate the QALYs gained by a new therapy. Therefore, some analysts have used bespoke ‘utility’ measures, willingness-to-pay estimates or discrete choice experiments in order to assess the various attributes of MDs. For example, Chancellor et al (2008) used bespoke utility scenarios to assess the benefits of inhaled insulin (using a device) as compared with injected insulin. They found a marginal increase in utility owing to the greater convenience offered by the inhaler.

In a systematic review of willingness-to-pay studies and discrete choice experiments of devices, Wilkinson and Drummond (2015) found 30 relevant empirical studies, covering technologies such as hearing aids, insulin delivery systems, Cochlear implants and endoscopy. They found that, in addition to clinical effectiveness, the studies assessed the value of attributes such as convenience/ease of use, length and/or frequency of treatment, visibility/size and method of treatment delivery.

These studies are useful in exploring which attributes of treatment are valued by patients and the trade-offs they make among the various attributes, but their use in resource allocation decisions is uncertain. Some health care decision-makers may not feel it appropriate to pay extra for benefits like increased convenience, unless this leads to an increase in treatment adherence and hence outcome. Also, many decision-makers value the standardization in methods offered by the generic preference-based measures, even if they fail to reflect some of the benefits of therapies.

2.2.4 Assessing organizational impact

As mentioned earlier, the potential for organizational impact is often mentioned in economic evaluations of MDs, but almost never explored quantitatively. Therefore, it is important that economic evaluations of MDs consider a wide range of costs; specifically, the costs of the infrastructure necessary to accommodate the device (eg catheter laboratories of hybrid operating rooms, the cost of training in the use of the new device and the creation of multidisciplinary teams). In addition, any impact of the new device on procedure costs should be quantified, by measuring any increase in the volume in procedures, or the conversion of procedures from an inpatient to an outpatient setting.

2.2.5 Accounting for dynamic pricing

Economic evaluations of MDs should anticipate future price changes by conducting a sensitivity analysis of device acquisition cost, or by identifying the threshold price at which the device is no longer cost-effective. For example, in the NICE re-appraisal of drug-eluting stents mentioned above (NICE, 2003), it was determined that DES would be cost-effective unless the price was more than £300 higher than the price of bare metal stents.

3. Managing the diffusion of MDs

Understanding the factors that lead to the adoption of MDs is especially important to health care decision makers because the regulatory barriers to market access are considerably lower for MDs than for pharmaceuticals. A number of studies have shown that the utilization of MDs varies substantially, both within and between countries (Valzania et al, 2015). In a recent study, Torbica et al (2017) used a panel data regression model to investigate the variation in implant rates for cardiac implantable electrical devices across 57 regions in 5EU countries. By analysing a total of 1,330,098 hospitalizations extracted from discharge databases in Austria, England, Germany, Italy and Slovenia from 2008 to 2012, they found implant rates were positively associated with higher levels of tertiary education among the labour force and the percentage of population over 74 years old. Regional per capita GDP and the number of implanting centres appeared to have no significant effect. There were also significant country effects in some of the analyses, suggesting that institutional factors play a role.

In another recent study, Hatz et al (2017) investigated the characteristics of organizations (ie hospitals) and individuals (ie clinicians) that are more inclined to adopt and utilize cardiovascular devices based on a comprehensive analysis of environmental, organizational, individual and technological factors. Seven random intercept hurdle models were estimated using data obtained from 1249 surveys completed by members of the European Society of Cardiology. They found that better (device) manufacturer support increased the adoption probability of new cardiac devices (defined by CE mark approval dates) and that budget pressure increased the adoption probability of older devices. They argue that the role of manufacturer support should be investigated in more detail in order to assess whether it functions as a substitute for medical evidence on new devices and to gain insights about its relationship with clinical and cost-effectiveness.

This goes to the heart of the matter. Namely, if wide variation is observed in the utilization of MDs across countries and regions, is the level of utilization in each jurisdiction 'appropriate' (defined in terms of clinical and cost-effectiveness), or are devices being overused in some jurisdictions and inappropriately restricted in others? Ideally one would investigate the link between the use of the technologies and the patient outcomes obtained, but this is not possible using routinely available data. Alternatively, it may be possible to investigate whether physicians' practices are in line with accepted clinical guidelines, although with a few notable exceptions (eg NICE's clinical guidelines in the UK; www.nice.org.uk) these do not consider cost-effectiveness.

Another approach would be to use economic evaluation to manage the diffusion of MDs, by conducting studies not only at market entry, but also in the post-market phase. In recent years there has been a growth in managed entry agreements that are collectively known as performance-based risk-sharing arrangements (PBRSAs). These schemes are known by different names, and take slightly different forms, in different jurisdictions, such as 'coverage with evidence development' in the US, 'field evaluations' in Canada or 'only with research' in the UK. The basic approach is that new technologies are reimbursed by the health care system on the condition that further research is conducted into the clinical and/or cost-effectiveness of the technology. This approach is particularly well-suited to MDs, since they are less well studied in clinical trials in the pre-market phase and that, because of the potential for a learning curve and incremental innovation, it is likely that clinical and cost-effectiveness in regular clinical practice will differ substantially from that observed in an experimental setting such as a controlled clinical trial.

Garrison et al (2013) reviewed PBRSAs existing in a number of countries. Under their working definition, a PBRSA exhibits the following key characteristics (i) there is a programme of data collection agreed between the manufacturer and the payer (ii) the data collection is typically initiated during the time period following regulatory approval (iii) the price, reimbursement and/or revenue for the product are linked to the outcome of the programme of data collection (iv) the data collection is intended to address uncertainty (eg the uncertainty could be about the efficacy or effectiveness in the treatment population tested in the pre-market phase, the efficacy or effectiveness in a broader population or in the long term, adverse events or adherence issues, whether health care providers' actual management of patients affects outcomes, the size and value of cost offsets, the number and attributes of the patients treated in practice) and (v) the arrangements provide a different distribution of risk between the payer and manufacturer than under the normal reimbursement arrangements.

Examples of PBRSA for procedures involving devices initiated in the US include those for positron emission tomography for cancer diagnosis, percutaneous transluminal angioplasty and stenting for secondary prevention of stroke and spinal cord stimulation for failed back surgery syndrome (Garrison et al, 2013). In another example of a PBRSA in Ontario (Canada), the generalizability of existing randomised controlled trials of drug-eluting stents (DES), which had been conducted in the US, was questioned. Therefore, a pragmatic registry of all patients receiving DES was established, in order to conduct a 'field evaluation'. Coverage was provided for the stents provided under the registry. It was found that DES was more effective only in patients at high risk of stenosis (eg those with diabetes, or particularly long or narrow lesions). Since this represented about 30% of the whole patient population, it was argued that this policy saved between $35-58 million, as compared with the potential uncontrolled adoption of DES (Levin et al, 2007; Goeree et al, 2009).

The key questions in using economic evaluation to manage the diffusion of MDs concern the timing of studies (eg before or after providing reimbursement) and the nature of the research (eg the parameters estimated and the methods used). Making a device available early can have benefits, as it allows patient access to potentially beneficial therapies. However, decisions about the use of the device when the evidence base is relatively immature can lead to clinical and financial risk. Also, once the device is generally available, the possibilities for further research to reduce uncertainty may be limited, either because manufacturers have fewer incentives to invest in further research, or because patients and clinicians are unwilling to participate in studies, especially RCTs.

Health economists have explored the use of value of information analysis to determine the balance between the benefits of early access to a technology and the value of additional evidence to resolve uncertainty (Claxton et al, 2012). In a recent paper, Rothery et al (2017) set out a framework for considering the issues relating to the uncertainty in the diffusion and use of MDs, considering the learning curve, incremental innovation, investment and irrecoverable costs and dynamic pricing. They use this framework to illustrate the circumstances under which it might be preferable to restrict access to the device until further research establishes its value, or preferable to allow access as long as further research is conducted.

Rothery et al (2017) also discuss the incentives for further research and how the value and cost of such research might be shared between the manufacturer and the health care sector. Drummond et al (2016) argue that incentives to manufacturers to conduct research need particular attention, give the lack of data exclusively mentioned earlier. In a situation where a PBRSA has been established for the first device to market, this might provide an opportunity for different manufacturers to share the costs of the research, as opposed to most of the costs being borne by the first to market. For example, 'fast follower' devices claiming substantial equivalence could be included in the scheme, thereby sharing the costs of the research and generating evidence on the comparative effectiveness of the different devices.

4. Concluding remarks

In this paper we have outlined the key challenges in conducting economic evaluations of MDs and how they might be tackled. However, many methodological and policy challenges remain and further research is required. First, approaches need to be developed to integrate regulatory and health technology assessment for MDs, so that future data needs can be anticipated when clinical studies are being planned. It is likely that the balance of data collection for the economic evaluation of

MDs will be in the post-marketing phase. Therefore, effort needs to be placed of the analysis of 'real-world' costs and effects of devices in regular clinical use. Linked to this, attention needs to given to the design of clinical registries, so that these will also be useful for conducting economic evaluations, as well as establishing the efficacy and safety of MDs.

Secondly, more study is required to explore the impact of the various distinctive characteristics of medical devices. In particular, it would also be useful to determine whether it is possible categorize devices in terms of their likely learning curve, according to the types of technology (diagnostic or therapeutic; implantable or non-implantable), the medical specialty, or device risk class. In additional, there should be more quantitative assessments of the organizational impact of MDs

Finally, further study is required of the various influences on the diffusion of devices, especially the motivational factors facing patients, clinicians and the institutions in which they work. In turn this could help improve the design and conduct of PBRSAs, which are likely to be central to the economic evaluation of MDs in the future.

Further reading

(a) *The specific characteristics of* MDs: the following two papers discuss the arguments for and against considering the specific characteristics of medical devices when conducting economic evaluations.

Drummond, M.F., Griffin, A., Tarricone, R. (2009) Economic evaluation for devices and drugs. Same or different? *Value in Health*; 12(4): 402-404.

(b) *The need to strengthen the regulation of* MDs: this paper discusses this issue in the context of the US and Europe.

Sorenson, C., Drummond, M.F. (2014). Improving medical device regulation: the US and Europe in perspective. *The Milbank Quarterly*; 92:1, 114-150. Doi: 10.1111/1468-0009.12043.

(c) *Challenges in the assessment of medical* devices:this journal supplement describes the results of the European Union MedtecHTA Project, a recent and comprehensive research effort in the economic evaluation of MDs.

Tarricone, R, Torbica, A, Drummond, MF, Schreyögg J. (2017). Assessment of medical devices: challenges and solutions. *Health Economics*; 26(Suppl. 1): 1-152.

References

Armeni P, Boscolo PR, Tarricone R, Capodanno D, Maggioni AP et al (2016). Real-world cost effectiveness of MitraClip combined with medical therapy versus medical therapy alone in patients with moderate of severe mitral regurgitation. *International Journal of Cardiology*; 209: 153-160.

Bell R, Pandanaboyana S, Hanif F, Shah N, Hidalgo E, Lodge JPA, Toogood G,

Prasad KR (2015). A cost effective analysis of a laparoscopic versus an open left lateral sectionectomy in a liver transplant unit. *HPB (Oxford);17(4):332-6.*

Burch J, Epstein D, Baba-Akbari et al (2008) Stapled haemorrhoidectomy (haemorrhoidopexy) for the treatment of haemorrhoids: a systematic review and economic evaluation. *Health Technology Assessment* ;12:1-214.

Chancellor J, Aballea S, Lawrence A et al (2008) preferences of patients with diabetes for inhaled versus injectable insulin regimens. *Phamacoeconomics*; 26: 217-234.

Ciani O, Wilcher B, van Giessan A, Taylor RS (2017) Linking the regulatory and reimbursement processes for medical devices: the need for integrated assessments. *Health Economics*; 26 (suppl 1): 13-29.

Claxton K, Palmer S, Longworth L et al (2012) Informing a decision framework for when NICE should recommend the use of technologies only in the context of an appropriately designed programme of evidence development. *Health Technology Assessment*; 16:1-103.

Drummond, M.F. (2013) Twenty years of using economic evaluations for drug reimbursement decisions. What has been achieved? *Journal of Health Politics, Policy and Law* 38(6): 1081-1102.

Drummond MF, Griffin A, Tarricone R (2009) Economic evaluation for devices and drugs – same or different? *Value in Health*; 12: 402-404.

Drummond MF, Tarricone R, Torbica A (2016) Incentivizing research into the effectiveness of medical devices. *European Journal of Health Economics*;doi 10.1007/s10198-016-0823-3.

Dyer MT, Goldsmith KA, Khan SN, Sharples LD, Freeman C, Hardy I, Buxton MJ, Schofield MF (2008) Clinical and cost-effectiveness analysis of an open label, single-centre, randomised trial of spinal cord stimulation (SCS) versus percutaneous myocardial laser revascularisation (PMR) in patients with refractory angina pectoris: The SPiRiT trial. *Trials*; 9: 1-11.

Food and Drug Administration (2016) Unsafe and ineffective devices approved in the EU that were not approved in the US. http://www.elsevierbi.com/~/media/Supporting%20Documents/The%20Gray%20Sheet/38/20/FDA_EU_Devices_Report.pdf. (Accessed 18 July 2016)

Fraser AG, Krucoff MC, Brindis RG, Komajda M, Smith SC Jr (2011) Commentary: International collaboration needed on device clinical trials. *British Medical Journal*; 342: d2952.

Garrison L, Towse A, Briggs A, de Pouvourville G, Greuger J, Mohr PE, Severens JL, Siviero P, Sleeper M (2013) Performance-based risk-sharing arrangements – Good practices for design, implementation, and evaluation: Report of the ISPOR Good Practices for Performance-Based Risk-Sharing Arrangements Task Force. *Value in Health*; 16: 703-719.

Goeree R, Bowen JM, Blackhouse G, Lazzam C, Cohen E, Chiu M, Hopkins R,

Tarride J-E, Tu, JV (2009) Economic evaluation of drug-eluting stents compared to bare metal stents using a large prospective study in Ontario. *International Journal of Technology Assessment in Health Care*; 25: 196-207.

Grieve R, Abrams K, Claxton K, Goldacre B, James N, Nicholl J, Parar M, Parker C, Sekhon S, Smeeth L, Spiegelhalter D, Sculpher M (2016) Cancer Drugs Fund requires further reform. *British Medical Journal*; 354 doi: http://dx.doi.org/10.1136/bmj.i5090

Guillou PJ, Quirke P, Thorpe et al (2005) Short-term endpoints of conventional versus laparoscopic-assisted surgery in patients with colorectal cancer (MRC CLASICC trial); multicentre, randomised controlled trial. *Lancet*; 365:1718-1726.

Hatz MHM, Schreyögg J, Torbice A, Oriani G, Blankart CRB (2017) Adoption decisions for medical devices in the field of cardiology: results from a European survey. *Health Economics*; 26 (suppl 1): 124-144.

Husereau, D., Drummond, M.F., Petrou, S., Carswell, C., Moher, D., Greenberg, D., *et al.* (2013). Consolidated health economic evaluation reporting standards (CHEERS) - explanation and elaboration: A report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value in Health,* 16(2): 231-250.

Klarman HE, Francis J, Rosenthal G (1968) Cost-effectiveness analysis applied to the treatment of chronic renal disease. *Medical Care*; 6: 48-54.

Levin L, Goeree R, Sikich N, Jorgensen B (2007). Establishing a comprehensive continuum from an evidentiary base to policy development for health technologies: The Ontario experience. *International Journal of Technology Assessment in Health Care*; 23:299-309.

Lin P-J, Cangelosi MJ, Lee DW, Neumann PJ (2013). Willingness to pay for diagnostic technologies: a review of the contingent valuation literature. *Value in Health*; 16: 797-805.

Margier J, Tchouda SD, Banihachemi JJ, Bosson JL, Plaweski S (2015) Computer-assisted navigation in ACL reconstruction is attractive but not yet cost efficient. *Knee Surg Sports Traumatol Arthrosc; 23(4):1026-34.*

National Institute of Health and Care Excellence (NICE) (2008) *Drug-eluting stents for the treatment of coronary artery disease*. Technology Appraisal 152; guidance:4.

Office of Technology Assessment (1978) *Policy implications of the computed tomography (CT) scanner*. Library of Congress Catalog Number 78-600078. www.ota.fas.org (Accessed 22 December 2016).

Prentice JC, Conlin PR, Gellad WF, Edelman D, Lee TA, Pizer SD. Capitalizing on prescribing pattern variation to compare medications for type 2 diabetes. *Value in Health* 2014; 17: 854–62.

Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF (2001). Statistical assessment of the learning curves of health technologies. *Health Technology Assessment*; 5(12): 1-104.

Rognoni C, Ciani O, Sommariva S, Tarricone R (2016). Real-world data for the

evaluation of transarterial radioembolization versus sorafenib in hepatocellular carcinoma: a cost-effectiveness analysis. *Value in Health*; http://dx.doi.org/10.1016/j.jval.2016.09.2397

Rome BN, Kramer DB, Kesselheim AS (2014) FDA approval of cardiac implantable electronic devices via original and supplement premarket approval pathways (2014) *Journal of the American Medical Association*; 311: 385-391.

Rothery C, Claxton K, Palmer S, Epstein D, Tarricone R, Sculpher M (2017) Characterising uncertainty in the assessment of medical devices and determining future research needs. *Health Economics*; 26 (suppl. 1): 109-123.

Schnell-Inderst P, Iglesias C, Arvandi M, Ciani O, Matteucci Gothe R, Peters j, blom AW, Taylor RS, Siebert U (2017) Synthesis of RCT and Observational data: a bias-adjusted evidence synthesis of total hip replacement. *Health Economics*; 26 (suppl. 1): 46-69.

Shuren J, Califf RM (2016) Need for a national evaluation system for health technology. *Journal of the American Medical Association*; doi:10.1001/jama.2016.8708.

Sorenson C, Drummond MF (2014) Improving medical device regulations: the United States and Europe in perspective. *Milbank Quaterly;* 92: 114-150.

Tarricone, R., Drummond, M.F. (2011) Challenges in the clinical and economic evaluation of medical devices: The case of transcatheter aortic valve implantation. *Journal of Medical Marketing* 11(3) 221-229.

Tarricone R, Torbica A, Drummond MF (2017a) Challenges in the assessment of medical devices: the MedtecHTA project. *Health Economics*; 26 (suppl.1): 5-12.

Tarricone R, Callea G, Ogorevc M, Prevolnik Rupel V (2017b) Improving the methods for the economic evaluation of medical devices. *Health Economics*: 26 (suppl. 1): 70-92 .

Tarricone R, Torbica A, Drummond MF (2017c) Key recommendations from the MedtecHTA project. *Health Economics*; 26 (suppl. 1): 145-152.

Tarricone, R., Ferrè, F., Torbica, A., Drummond, M.F. (2014) Generating appropriate clinical data for value assessment of medical devices: what role does regulation pay? *Expert Review of Pharmacoeconomics & Outcomes Research,* 14:5, 707-718.

Taylor RS, Iglesias CP (2009) Assessing the clinical and cost-effectiveness of medical devices – are they that different? *Value in Health*; 12: 404-406.

Torbica A, Banks H, Valzania C, Boriani G, Fattore G (2017). Investigating regional variation of cardiac implantable electrical device implant rates in European healthcare systems: what drives differences? *Health Economics*; 26 (suppl. 1): 30-45.

Valzania C, Torbica A, Tarricone R, Leyva F, Boriani G (2015) Implant rates of cardiac implantable electrical devices in Europe: a systematic literature review. *Health Policy*; doi:1.1016/j.healthpol.2015.11.001.

van Gestel R, De Graeve D, Vrints C, Rodrigus I, Bosmans B (2013). Hospitalization costs of TAVI in one Belgian university hospital. *Acta Cardiologica*; 68: 263-270.

Varabyova Y, Blankart CR, Schreyögg J (2017) The role of learning in health technology assessments: an empirical assessment of endovascular aneurysm repairs in German hospitals. *Health Economics*; 26 (suppl.1): 93-108.

Wilkinson, G., Drummond, M.F**.** (2015). Alternative approaches for assessing the socioeconomic benefits of medical devices: a systematic review. *Expert Review of Medical Devices, 12*(5), 629-648. doi: 10.1586/17434440.2015.1080118.