New Metrics for Economic Evaluation in the Presence of Heterogeneity: Focusing on Evaluating Policy Alternatives Rather than Treatment Alternatives

David D. Kim, PhD, Anirban Basu, PhD

Background. *Cost-effectiveness analysis (CEA) methods fail to acknowledge that where cost-effectiveness differs across subgroups, there may be differential adoption of technology. Also, current CEA methods are not amenable to incorporating the impact of policy alternatives that potentially influence the adoption behavior. Unless CEA methods are extended to allow for a comparison of policies rather than simply treatments, their usefulness to decision makers may be limited.* ***Methods.*** *We conceptualize new metrics, which estimate the realized value of technology from policy alternatives, through introducing subgroup-specific adoption parameters into existing metrics, incremental cost-effectiveness ratios (ICERs) and Incremental Net Monetary Benefits (NMBs). We also provide the Loss with respect to Efficient Diffusion (LED) metrics, which link with existing value of information metrics but take a policy evaluation perspective. We illustrate these metrics using policies on treatment with combination therapy with a statin plus a fibrate v. statin monotherapy for patients with diabetes and mixed dyslipidemia.* ***Results.*** *Under the traditional approach, the population-level ICER of combination v. monotherapy was $46,000/QALY. However, after accounting for differential rates of adoption of the combination therapy (7.2% among males and 4.3% among females), the modified ICER was $41,733/QALY, due to the higher rate of adoption in the more cost-effective subgroup (male). The LED metrics showed that an education program to increase the uptake of combination therapy among males would provide the largest economic returns due to the significant underutilization of the combination therapy among males under the current policy.* ***Conclusion.*** *This framework may have the potential to improve the decision-making process by producing metrics that are better aligned with the specific policy decisions under consideration for a specific technology.* ***Key words:*** *economic evaluation; policy; technology diffusion; heterogeneity; cost-effectiveness.* ***(Med Decis Making XXXX;XX: xx–xx)***

Received 18 April 2016 from Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA, USA (DDK); Pharmaceutical Outcomes Research and Policy Program, Department of Pharmacy, and the Departments of Health Services and Economics University of Washington, Seattle, WA, USA (AB). Revision accepted for publication 5 March 2017.

Supplementary material for this article is available on the Medical Decision Making Web site at http://journals.sagepub.com/home/mdm.

Address correspondence to David D. Kim, PhD, Center for the Evaluation of Value and Risk in Health, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, 800 Washington St. Box 63, Boston, MA 02155, USA; Phone: (617) 636-5769, Fax: (617) 636-8628, e-mail: dkim3@tuftsmedicalcenter.org.

DOI: 10.1177/0272989X17702379

Since its development, cost-effectiveness analysis (CEA) has been one of the most prevalent methods in economic evaluations.[1,2] CEA has followed the notion that, if a technology is deemed cost-effective, it should be made available for use in the healthcare system.[3,4] However, determination of the incremental cost-effectiveness ratio (ICER) is usually based on intention-to-treat effects from clinical trials, which rely on adoption rates within the trial. Certainly, under a homogeneous treatment-effects assumption, differential adoption of this technology across subgroups would yield the same population-level ICER. When this stringent assumption is relaxed, however, a CEA result becomes limited because it does not incorporate the extent to which medical technology is used in practice and how this extent of use is likely to change over time (i.e., technology diffusion).[5,6]

From an evaluation perspective, such practice of CEA is possibly misleading for 2 reasons. First, it fails to acknowledge evidence-driven differential adoption across subgroups that can inherently change the overall value of medical technology. Second, it does not allow for evaluating the impact of policy alternatives that potentially influence the adoption behavior. Such policies may include different levels of insurance coverage or the implementation of practice guidelines to improve (or discourage) the use of (non-) cost-effective technology in practice. In essence, we argue that the incremental value of technology should not be reflected by a single metric, such as an ICER or a net benefit metric, but rather tied to the particular policy that would change the adoption of this technology.

Within this context, we first set up a conceptual framework that could integrate heterogeneity in costs and effectiveness across subgroups with their corresponding adoption patterns into the existing metrics in Section 2. Then, we show how this new framework can be related to the existing value of information framework in Section 3, followed by providing overarching metrics, "Loss with respect to Efficient Diffusion (LED)," which can be used to prioritize policy alternatives in Section 4.

This framework extends previous research on heterogeneity in economic evaluation. For example, Phelps[7] introduced the idea that variations in cost-effectiveness ratios can be driven by heterogeneity, after which Coyle and others[8] quantified the potential health gains facilitated by making different decisions for different subgroups. Basu and Meltzer[9] extended this concept to decisions at the individual level to estimate the potential value of providing information on patient preference to make individualized treatment decisions. Espinoza and others[10] showed how variability (observed heterogeneity) and uncertainty interacts through the static value of heterogeneity and the dynamic value of heterogeneity. In this paper, we extended the existing literature by introducing policy-specific metrics, rather than treatment-specific metrics, by incorporating adoption behavior induced by such policies.

We provide an empirical illustration on how this new framework could be applied in a real-world example in Section 5 and conclude with a discussion in Section 6.

CONCEPTUAL FRAMEWORK FOR POLICY EVALUATIONS

Let us suppose that a social insurer aims to maximize a health outcome (e.g., quality-adjusted life years, QALYs) given a fixed budget. The social insurer faces a policy decision about whether to provide coverage for a new treatment A for treating a chronic illness v. the standard of care, as denoted by B. Traditionally, one would conduct a CEA comparing treatment A v. B and estimate an ICER. If the ICER is less than a cost-effectiveness threshold, λ, it indicates that paying for treatment A will provide good value and vice versa. (Based on the constrained optimization framework, λ can represent the rate at which QALYs would be forgone elsewhere in the healthcare sector if the new treatment were paid for, given the fixed budget.) While these ideas pervade in the cost-effectiveness literature, there exists the implicit assumption behind the interpretation of ICER, which assumes that, if covered, treatment A would be immediately used/adopted by the same proportion of patients as in the supporting RCT.

However, there is increasing recognition that heterogeneity in costs, effectiveness, and cost-effectiveness should be considered.[5,11,12] The role of evidence on heterogeneity in influencing individual decision making is relatively straightforward when decision making is not centralized. In this setting, the expected value of individualized care (EVIC) may be calculated to establish the value of generating evidence to improve individual decision making.[9] However, the role of heterogeneity in population-level (centralized) policy decision making is less clear. Some have argued that subgroup analyses should be carried out in CEA when the goal is to provide differential coverage across these subgroups.[13,14] While new research has delved into the methodological issues of estimating and presenting heterogeneity in cost-effectiveness,[10] little attention has been given to how such information can be best used to inform population-level decision making.

To illustrate the issues raised by heterogeneity in cost-effectiveness, we assume that there are 3 subgroups (j = 1, 2, 3) indexed by an easily observable characteristic, such as age. Let these estimates of cost-effectiveness represented by ICERs be:

$$ICER_{j,AB} = \frac{E(C_{A,j}) - E(C_{B,j})}{E(Q_{A,j}) - E(C_{B,j})} \quad \text{j}=1,2,3 \qquad (1)$$

Let $ICER_{1,AB}$ and $ICER_{2,AB}$ be both greater than λ, and $ICER_{3,AB}$ be less than λ. (i.e., $ICER_{1\&2,AB} > \lambda > ICER_{3,AB}$). That is, compared to the threshold λ, treatment A would generate sufficient value for subgroup 3 but not for subgroups 1 and 2. Despite this evidence of heterogeneity, the traditional ICER that

is used to inform a uniform coverage decision for treatment A, and is assumed to be $> \lambda$, can also be written as:

$$ICER = \frac{E(C_A) - E(C_B)}{E(Q_A) - E(Q_B)} = \frac{\sum_j \{P_j \cdot [E(C_{A,j}) - E(C_{B,j})]\}}{\sum_j \{P_j \cdot [E(Q_{A,j}) - E(C_{B,j})]\}} > \lambda \quad (2)$$

In the above formula, the size of these subgroups is given as $P_j \leq 1$ for $j = 1, 2, 3$. It is assumed that when treatment A is covered, it would be adopted by all patients in all 3 subgroups, despite evidence that it is not cost-effective for the first 2 subgroups.

However, a modified ICER can be presented by rewriting the traditional ICER in the scenario that this heterogeneity could lead to differential adoption as:

$$ICER_{NEW} = \frac{\sum_j \{P_j \cdot D_j \cdot [E(C_{A,j}) - E(C_{B,j})]\}}{\sum_j \{P_j \cdot D_j \cdot [E(Q_{A,j}) - E(C_{B,j})]\}} = \frac{\sum_j \{P_j \cdot D_j \cdot E(\Delta C_j)\}}{\sum_j \{P_j \cdot D_j \cdot E(\Delta Q_j)\}} \quad (3)$$

In the above formula, D_j represents the rate of adoption of treatment A in the population subgroup j, and can be estimated as following:

$$D_j = \frac{\sum_{i=1}^{N_j} I(D_i)}{N_j} \quad (4)$$

where N_j is the size of the population subgroup j who are receiving, or are expected to receive, one of the comparators, and $I(D_i)$ indicates whether an individual patient in the specific subgroup received treatment A when given full coverage to the treatment. This formulation can be easily adapted to capture adoption over time, but we keep that implicit for simplicity of illustration. If the rate of adoption of the technology is same in all 3 subgroups, we get back the original ICER as shown in eq. (2). However, if the adoption rates are different, as would be expected based on between-subgroup differences in cost-effectiveness, the modified ICER would be a better reflection of the realized value of this technology in practice.

One can extend this concept to allow for evaluating different policies. For example, the ICERs from an insurer perspective could be very different if the underlying policies were to cover A at 50% cost-sharing v. no coverage, holding coverage for treatment B constant. This implies that the traditional ICER, as expressed in eq. (1), cannot readily be used to inform more nuanced forms of coverage policies that many public and private payers contemplate, whereas the modified ICER in eq. (3) can help address this issue. Furthermore, if the social insurer takes a narrower perspective as a payer, the ICER under a policy k, which only accounts for a fraction, f_k, of the incremental costs (i.e., the rest being borne by the patient as cost-sharing), is given as:

$$ICER_{policy_k} = \frac{\sum_j \{P_j \cdot D_{jk}(f_k) \cdot E(\Delta C_j) \cdot f_k\}}{\sum_j \{P_j \cdot D_{jk}(f_k) \cdot E(\Delta Q_j)\}} \quad (5)$$

where $D_{jk}(f_k)$ is the expected adoption of treatment A in subgroup j under Policy k. Note that the numerator includes not only a fraction of the incremental costs borne by the payer but also the adoption probability is tempered with cost-sharing in line with price elasticities of demand. For simplicity, we assume that f_k applies to incremental costs, which assumes f_k applies to both treatment A and B. To relax this assumption, f_k can be treated as a treatment-specific parameter, $f_{k,j}$. Also, CEA generally includes not only treatment costs but also other monetary consequences in the numerator as well. One can also separate out the treatment costs from the other pecuniary consequences of treatment, and apply the cost sharing to the former but not the latter.

From a healthcare sector perspective, which accounts for all cost of care irrespective of who bears them, the full incremental costs will be accounted for but the adoption rate would still be driven by cost-sharing. Thus, from the perspective of the healthcare sector:

$$ICER_{policy_k} = \frac{\sum_j \{P_j \cdot D_{jk}(f_k) \cdot E(\Delta C_j)\}}{\sum_j \{P_j \cdot D_{jk}(f_k) \cdot E(\Delta Q_j)\}} \quad (6)$$

This new ICER formulation, as expressed in eq. (5) or eq. (6), is better-suited for informing policy making across alternative policy options in the presence of heterogeneity according to relative effectiveness or costs, patient preference, or patient choice.[15] Standard decision-making criteria, where the ICER is compared to the cost-effectiveness threshold, could be applied here. Note that the biggest difference between this and the traditional ICER is that the traditional formulation provides only one ICER for comparing treatment A to treatment B for each subgroup. Under the new formulation, the ICER differs across policy decisions, since these policy decisions could have differential effects on the rate of

adoption of treatment A. This new metric allows the adoption rates to be endogenous to the policy decision itself, and consequently, generates a more policy decision-relevant estimate of relative value.

Also, we argue that net benefit framework may be better suited to make these multiple comparisons from a policy evaluation perspective (see Appendix A1 for more details). In this setting, the net monetary benefits (NMB) under a policy k, which could have implications for cost-sharing and/or adoption of any treatment t within subgroup j, would be given as:

$$NMB_{policy_k} = \sum_j \sum_t \left[P_j \cdot D_{jkt}(f_k) \cdot \left(E(Q_{jt}) \cdot \lambda - E(C_{jt})\right)\right] \quad (7)$$

Naturally, with multiple policy options, equation (7) would be better suited to compare these policy options. In fact, one can also account for differential costs of policy implementation (M_K) in this framework and express eq. (7) as:

$$NMB_{policy_k} = \sum_j \sum_t \{P_j \cdot D_{jkt}(f_k) \cdot \left[E(Q_{jt}) \cdot \lambda - E(C_{jt})\right]\} - M_K \quad (8)$$

The policy that maximizes NMB would be deemed as the optimal policy given λ. Therefore, the policy-specific ICERs or NMBs could be used to make policy decisions across a broad range of the decision space. Evaluating medical technology through the lens of specific policies changes the calculus for expected realized value from that technology. Consequently, it should also affect the expected value of future research on this technology.

In the next section, we develop 2 concepts termed ''Efficient diffusion'' and ''Loss with respect to efficient diffusion'' in the following sections. Most of these concepts largely build on the existing framework developed by Coyle and others,[8] Basu and others,[9] Fenwick and others,[16] and Espinoza and others;[10] however, we expanded the existing framework from a policy evaluation perspective.

EFFICIENT DIFFUSION UNDER UNCERTAINTY V. CERTAINTY

Efficient Diffusion or Stratification

Efficient diffusion represents an ideal phenomenon where only those in the cost-effective subgroup(s) fully adopt the new treatment and those in the other subgroup(s) do not elect to adopt the new treatment. This is the same phenomenon laid out by Coyle and others[8] and Espinoza and others[10] around stratification or stratified care. Throughout this paper, we use the term ''diffusion'' to highlight the fact that such ideal stratification does not occur overnight and the diffusion patterns can be driven by policy alternatives.

If efficient diffusion can be achieved, then there is no need to restrict coverage of treatment A for everyone, as only those who would gain more than the opportunity cost of consuming treatment A would use treatment A. It produces the maximal value from treatment A given current evidence.[8] Thus, efficient diffusion implies that all technologies should be covered by insurance as long as those technologies are used in an efficient manner, a staple concept in the first-best solution of health insurance markets.[17,18] In the presence of multiple treatments (t = 1,2,..T), let efficient diffusion for any treatment t within a subgroup j be characterized using a binary indicator D^*_{jt}, depending on whether that treatment is cost-effective in that subgroup. (D* still represents the same conceptual adoption rate as D in the previous section. The only difference is that D* represents a theoretical and ideal phenomenon where the adoption of technology A is perfectly diffused to a specific subgroup or perfectly not adopted by the group. The asterisk highlights this ideal situation. D* can only take the value of either 1 or 0 for a subgroup, whereas D can be a continuous value between 0 to 1.)

There could be 2 types of efficient diffusion—one under current information, which includes the inherent uncertainty present in this evidence, and the other under perfect information, where all uncertainties are resolved. We now describe these 2 scenarios.

Efficient Diffusion under Current Information (EDCI)

Under current information, efficient diffusion represents perfect implementation (i.e., subgroup-level perfect adoption of technologies) based on what we currently know about the cost-effectiveness of new medical technology. It also implies that, given current evidence, if $NMB_j < 0, \forall_j$, then it is sufficient to forgo coverage decisions for new treatment A, as it is not cost-effective for any subgroup.

Under our stylized example of 2 treatments and 3 subgroups, efficient diffusion would imply that, even with full coverage of treatment A for all subgroups, only those in the third subgroup fully adopt treatment A (i.e., $D^*_{3,A}=1$ and $D^*_{3,B}=0$), whereas those in the first and second subgroups entirely reject treatment A (i.e., $D^*_{1\&2,A}=0$ and $D^*_{1\&2,B}=1$). Under this scenario, the NMB arising out of EDCI is given as:

$$NMB_{EDCI}=P_1\cdot E[(Q_{1B}\cdot\lambda-C_{1B})]+P_2\cdot E[(Q_{2B}\cdot\lambda-C_{2B})]+P_3\cdot E[(Q_{3A}\cdot\lambda-C_{3A})] \quad (9)$$

In a more generalized formula, the NMB under EDCI can be expressed as:

$$NMB_{EDCI}=\sum_j\sum_t\left\{P_j\cdot D^*_{jt}\cdot\left[E(Q_{jt})\cdot\lambda-E(C_{jt})\right]\right\}=\sum_j P_j\cdot max_t\left[E(Q_{jt}\cdot\lambda-C_{jt})\right] \quad (10)$$

Equation (10) is equivalent to the metrics previously described as the total net benefits by Coyle and others[8] and Espinoza and others.[10] It also represents a weighted average of maximum expected net benefit across subgroups, which is the second term of the traditional expected value of perfect information (EVPI) expression.[19] Note that this metric in eq. (10) is not policy-specific, as it represents the best one could do under current information through the perfect implementation of current evidence.

Efficient Diffusion under Perfect Information (EDPI)

Given the existing uncertainty in NMBs under current information, there is a possibility that even perfect implementation of current evidence would incur some losses. This is tied to the fact that whatever decision is deemed optimal under current information, it may be suboptimal when uncertainty is resolved. That is, research generating more precise estimates of existing evidence could induce more appropriate adoption of the new treatment across subgroups.

With the stylized example, current evidence suggests that new treatment A is cost-effective in subgroup 3 but not in subgroup 1 or 2. However, there is uncertainty associated with the incremental net monetary benefit (INMB) between the new treatment A and standard of care B for each subgroup. For example, let us state that the uncertainty is least for subgroup 3, and the results favor treatment A, but there is uncertainty nevertheless. To account for the existing uncertainty, we generate the cumulative distribution function for the INMB between A and B for each subgroup as $F_{INMB_j\cdot AB}(x)$ = $\Pr(INMB_{j,AB}\le x)$, given λ. Under this scenario, the probability that future research could change the current decision on the efficient use of a new treatment A is given as:

For Subgroup 1, Current Decision: Use B; Pr(A is optimal) = 1 -$F_{INMB_{1,AB}}$ (0)

For Subgroup 2, Current Decision: Use B; Pr(A is optimal) = 1 -$F_{INMB_{2,AB}}$ (0)

For Subgroup 3, Current Decision: Use A; Pr(B is optimal) = $F_{INMB_{3,AB}}$ (0)

Therefore, efficient diffusion that accounts for both today's evidence and the expected value of future research is given as:

$$\begin{aligned}NMB_{EDPI}=&P_1\cdot\{(Q_{1B}\cdot\lambda-C_{1B})+\left[1-F_{INMB_{1,AB}}(0)\right]\cdot \mathrm{E}(INMB_{1,AB}|INMB_{1,AB}>0)\}\\&+P_2\cdot\{(Q_{2B}\cdot\lambda-C_{2B})+\left[1-F_{INMB_{2,AB}}(0)\right]\cdot \mathrm{E}(INMB_{2,AB}|INMB_{2,AB}>0)\}\\&+P_3\cdot\{(Q_{3B}\cdot\lambda-C_{3B})+\left[1-F_{INMB_{3,AB}}(0)\right]\cdot \mathrm{E}(INMB_{3,AB}|INMB_{3,AB}>0)\}\\=&\sum_j P_j\cdot\left\{\begin{array}{l}\left[1-F_{INMB_{j,AB}}(0)\right]\cdot E\left(NMB_{j,A}|INMB_{j,AB}>0\right)\\+\left[F_{INMB_{j,AB}}(0)\right]\cdot E\left(NMB_{j,B}|INMB_{j,AB}\le 0\right)\end{array}\right\}\end{aligned} \quad (11)$$

Equation (11) indicates that, for each subgroup, there is some probability that the current decision to promote full adoption of A v. B as part of efficient diffusion may be wrong. And if we completely resolve this uncertainty, then we truly know which treatment to adopt for each subgroup. Thus, the expected NMB from efficient diffusion under perfect information (EDPI) would be an expectation over treatment-specific maximum NMBs when uncertainty is completely resolved. Note that this represents the weighted average of expected maximum net benefit across subgroups, which is the first term of the traditional EVPI expression; albeit, at the subgroup-specific level. This concept is also previously recognized by Espinoza and others[10] but, in our framework, it can be expressed in general as:

$$NMB_{EDPI}=\sum_j P_j\cdot E\left[max_t\left(Q_{jt}\cdot\lambda-C_{jt}\right)\right] \quad (12)$$

LOSS WITH RESPECT TO EFFICIENT DIFFUSION (LED) METRICS

To tie the existing value of information framework with the new metrics that we are proposing that incorporate policy-driven rates of selective adoption, we consider 3 different LED metrics. LED metrics estimate how much the realized value of policy alternatives deviates from the ideal scenario, called efficient diffusion or stratification (i.e., expressed in positive terms, such that the higher the value, the higher the loss, and the less attractive the policy). Figure 1 provides a graphic representation of new metrics along with the summary of these metrics.

$$LED_{EVPI} = NMB_{EDPI} - NMB_{EDCI} \quad (13)$$

$$LED_{Policy_K} = NMB_{EDCI} - NMB_{Policy_K} \quad (14)$$

$$LED^*_{Policy_K} = NMB_{EDPI} - NMB_{Policy_K} \quad (15)$$

LED_{EVPI} expresses the loss in expected value from efficient diffusion with perfect information (i.e., decision certainty) v. efficient diffusion with current information (i.e., decision uncertainty). Following eq. (10) and eq. (12), LED_{EVPI} is identical to traditional EVPI metric:[19]

$$LED_{EVPI} = \sum\nolimits_j P_j \cdot E\left[max_t\left(Q_{jt} \cdot \lambda - C_{jt}\right)\right] - \sum\nolimits_j P_j \cdot max_t\left[E(Q_{jt} \cdot \lambda - C_{jt})\right] \quad (16)$$

Here, it is interpreted as the maximum value of future research that would resolve uncertainty in each of the subgroups. This concept has also been presented as the dynamic value of heterogeneity by Espinoza and others.[10] This value is a comparison of two hypothetical scenarios involving efficient diffusions. The true value of future research could be different from what is typically expressed in EVPI calculations, and could depend on the current policy in place. We will now develop a second LED metric to exemplify this point.

LED_{Policy_K} expresses the expected incremental loss under a specific policy k with respect to NMB_{EDCI}, which is given as:

$$LED_{Policy_K} = \sum\nolimits_j \sum\nolimits_t \left\{P_j \cdot \left[D^*_{jt} - D_{jkt}(f_{kt})\right] \cdot \left[E(Q_{jt}) \cdot \lambda - E(C_{jt})\right]\right\} \quad (17)$$

In this formulation, irrespective of the value of f_{kt}, $[D^*_{jt} - D_{jkt}(f_{kt})] \leq 0$ in subgroups where $D^*_{jt} = 0$ (indicating over-adoption in these subgroups) while $\left[D^*_{jt} - D_{jkt}(f_{kt})\right] \geq 0$ in subgroup where $D^*_{jt} = 1$ (indicating under-adoption in these subgroups). The optimal policy decision, therefore, is based on choosing the one that minimizes this loss given the threshold λ and current evidence, as long as the 'status quo' (do-nothing) policy is included in these comparisons.

$$Policy_{K^*} = \underset{k}{\operatorname{argmin}}\, LED_{Policy_k} = \underset{k}{\operatorname{argmax}}\, NMB(Policy_k) \quad (18)$$

The utility of the LED_{Policy_K} metric mirrors the uses of the value of information metrics to prioritize implementation or research investments. For the optimal policy k^*, $LED_{Policy_{k^*}}$ represent the unrealized value due to over- and under-adoption of technologies in different subgroups driven by policy k^*, given current evidence. This loss may be caused by either imperfect implementation (i.e., subgroup-level suboptimal adoption due to imperfect translation of current evidence into practice) or imperfect information (i.e., suboptimal adoption due to too much uncertainty in existing information to make an optimal decision) or both. Therefore, depending on understanding the rationale for suboptimal adoption patterns, one could determine future investments on either the better implementation of current evidence or the generation of more precise estimates through research to capture this unrealized value (Appendix A2 for more details).

Finally, $LED^*_{Policy_K}$ would represent the expected loss relative to the expected value of perfection (i.e., perfect implementation with perfect information) under the policy k.[16] That is,

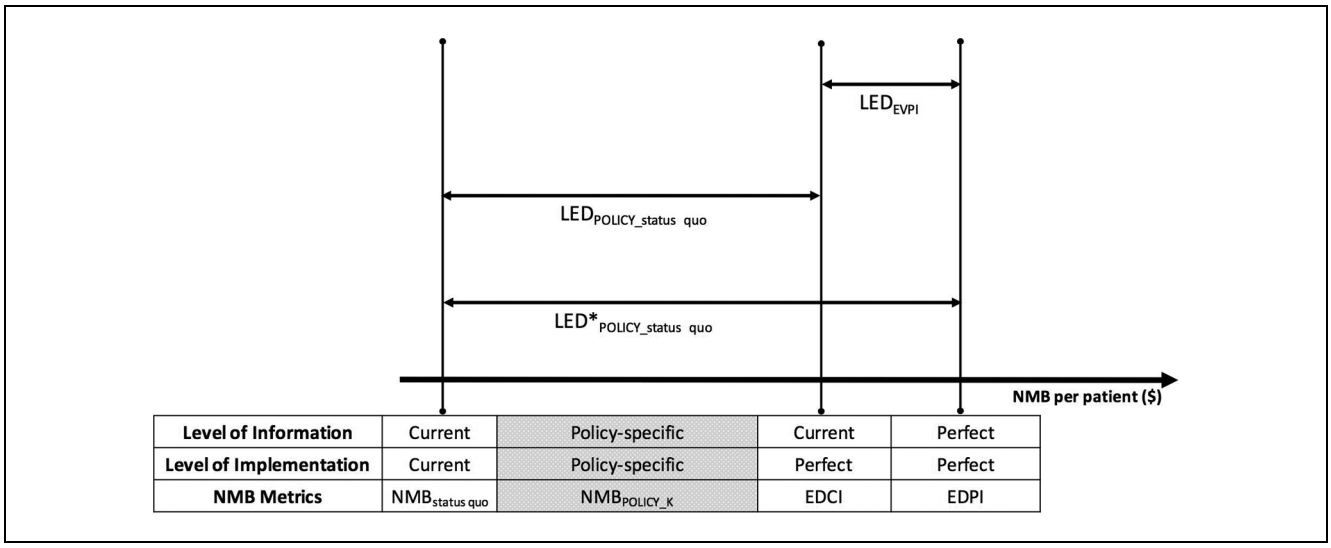
$$LED^*_{Policy_K} = LED_{EVPI} - LED_{Policy_K} \quad (19)$$

In the following section, we present evidence to support our novel metrics on the selective adoption of new technology under heterogeneous information. We then illustrate how our NMB_{Policy_K} and LED metrics would produce different suggestions for an optimal policy where we have accounted for such subgroup-specific adoption rates.

EMPIRICAL EXAMPLE

Background

The primary goal of this empirical example is to illustrate how this new framework could be applied in a real-world example. As an illustration, we



Level of Information	Current	Policy-specific	Current	Perfect
Level of Implementation	Current	Policy-specific	Perfect	Perfect
NMB Metrics	$NMB_{status\ quo}$	NMB_{POLICY_K}	EDCI	EDPI

Figure 1 Graphic representation of new metrics with summary of new metrics and equivalent metric/interpretation. LED^*_{POLICY}*, expected incremental loss under a specific policy* k *relative to the Expected Value of Perfection (EVP) (i.e., maximum value of future research and improved implementation);* LED_{EVPI}*, Expected Value of Perfect Information (EVPI) (i.e., maximum value of future research that would resolve all uncertainty);* LED_{POLICY}*, expected incremental loss under a specific policy* k *relative to the EVPIM (i.e., maximum value of improved implementation; there is a need to understand the rationale of suboptimal adoption patterns);* NMB_{EDCI}*, Expected Value of Perfect Implementation (EVPIM) (i.e., perfect implementation based on current evidence; suboptimal decision is possible);* NMB_{EDPI}*, Expected Value of Perfection (EVP) (i.e., perfect implementation with no decision uncertainty).*

chose the adoption of combination lipid therapy v. monotherapy among males and females with type 2 diabetes mellitus (T2DM) and mixed dyslipidemia based on 3 major criteria: 1) clinical evidence of heterogeneous treatment effects by subgroups based on a previous clinical trial (ACCORD study);[20] 2) subgroup-specific CEA results from a prior economic evaluation study;[21] and 3) the availability of datasets to estimate the rates of adoption of the intervention. Using the framework and metrics suggested, we evaluate the economic returns from policy alternatives for combination therapy with a statin plus a fibrate compared with statin monotherapy in this population among privately insured patients. Appendix A3 provides greater details on the background, estimation process, and data analysis.

Methods

We incorporated the clinical evidence presented by the ACCORD study and the reported subgroup–specific ICER estimates of combination therapy with a statin plus a fibrate (denoted as treatment A) compared with statin monotherapy (denoted as the standard of care, treatment B) amongst males and females (j =2).[21] We estimated subgroup-specific rates of adoption and the size of the relative population using the Truven Health Marketscan databases.[22] Then, we took the observed adoption behavior to address policy questions around coverage for the combination therapy.

We start by comparing the traditional overall ICER and INMB estimate in this population to that under the current status-quo policy, defined below, based on observed adoption rates, from a healthcare sector perspective. Also, we consider the corresponding ICERs and INMBs under alternative coverage policies. For the sake of illustration, we assume that the cost-effective threshold (CET) is $45,000/QALY, which corresponds to the 20,000-30,000/QALY threshold used by National Institute for Health and Care Excellence (NICE) in England and Wales. Specifically, we:

1. Report the traditional ICER and INMB following eq. (2), which is just the weighted average of subgroup-specific ICERs or INMBs, respectively. We use the estimated subgroup sizes from the Marketscan population to provide weighted estimates.
2. Report the modified ICER and INMB following eq. (3), where the differential adoption rate of the

combination therapy was also incorporated. However, this modified ICER or INMB was assumed to reflect a "status-quo policy" (#1), where both the combination therapy and statin monotherapy were offered at a 20% coinsurance rate (i.e., f_k = 0.80), as reflective of the average health plan coverage rate in the Marketscan database.

3. Report the modified ICER and INMB following eq. (6), dubbed as "Policy #2", where full coverage of the combination therapy would have been provided (i.e., f_k = 1.00), while coverage of the statin monotherapy remained unchanged at 20% coinsurance rate. For simplicity, we assume the same price elasticity of demand for a prescription fill for both subgroups. Average estimates of price-elasticity of the probability of fill were obtained from a recent study on value-based insurance design and were estimated to be -0.26.[23]
4. A "hypothetical policy (#3)" scenario, where full coverage of the combination therapy would have been provided (i.e., f_k = 1.00), while holding coverage of the monotherapy constant, and an implementation action (through an education outreach program) increased adoption of the combination therapy by 200% among males and decreased adoption of combination therapy by 100% among females. Although not included here, in a real application of such an approach, the cost of such a program should also be accounted for.

We also compare how the LED metrics vary under each of these scenarios. First, we aimed to incorporate uncertainty around the cost-effectiveness estimates measured for each subgroup (male and female). Unfortunately, the original CEA[21] did not report uncertainty estimates. Thus, for illustration, we simulated the standard errors to be 10% and 75% of the mean QALY and cost, respectively. We then generated a distribution of the expected INMB of the combination therapy compared to the statin monotherapy, based on 100,000 simulated values, and estimated the LED metrics using equations (13), (14) and (15).

Results

Traditional v. Modified Metrics under Alternative Policy Scenarios

Table 1 presents the results comparing traditional v. modified ICER and NMBs under alternative policy scenarios. Under the traditional approach, the ICER of combination therapy over statin monotherapy is estimated to be $46,000/QALY. Under the pre-specified threshold ($45,000/QALY), combination therapy would not be deemed cost-effective. Consequently, statin monotherapy should be adopted, and it produces a per-patient NMB of $315,869.

Policy 1: Status-quo policy. Under the status-quo policy of providing coverage for the combination therapy with a coinsurance rate of 20%, the use of this combination therapy is estimated to be 7.2% among the males and 4.3% among the females in the Marketscan database. Hence, in line with evidence, we empirically observe that males adopt the combination therapy at a higher rate than females given that the benefits to males are higher than to females. When these differential adoption rates were incorporated, the estimated ICER is $41,733/QALY, and the combination therapy would be deemed cost-effective. The per-patient NMB in the population under the status quo policy is $315,910.

Policy 2: Full coverage of the combination therapy. Under Policy 2, if the coverage of the combination therapy would have been increased to 100% (i.e., coinsurance rate = 0%), the share of use of the combination therapy is estimated to be 7.5% among males and 4.5% among females, which is not that much different to the status-quo policy. This is primarily because—as widely discussed following the RAND health insurance experiment—pharmaceuticals are very price inelastic;[24,25] in addition to the relatively low baseline adoption rates under the status-quo policy. Consequently, the ICER for combination therapy under Policy 2 is $41,766/QALY, and the per-patient NMB in the population under the Policy #2 is $315,911.

Policy 3: Full coverage of the combination therapy and Outreach Program. Finally, under the hypothetical policy, where not only combination therapy is fully covered, an educational outreach program is expected to increase adoption of combination therapy among males to 23% and to decrease in females to 2.3%, the ICER would be $34,848/QALY, and the NMB per patient would be $316,214. Obviously, one should account for the costs of delivering such an outreach program. For example, if the program costs about $200 per patient, the NMB decreases to $316,014, but still is the best policy among all the alternatives considered.

Table 1 Illustration of Traditional and Modified ICERs and INMBs under a Health-Care Sector Perspective

Parameter	Males	Females	Overall	Source / Notes
Total costs per patient under Statin + Fibrate, $	$107,021	$107,023	-	Based on Sorenson 2009[21]
Total costs per patient under Statin Only, $	$98,131	$98,131	-	Based on Sorenson 2009
Total incremental costs per patient, $	$8,890	$8,892	-	Based on Sorenson 2009
Total QALYs per patient under Statin + Fibrate, $	9.468	9.308	-	Based on Sorenson 2009
Total QALYs per patient under Statin Only, $	9.200	9.200	-	Based on Sorenson 2009
Total incremental QALYs per patient, $	0.268	0.108	-	Based on Sorenson 2009
Subgroup-specific ICER	$33,130/QALY	$82,562/QALY	-	
Subgroup-specific INMB[a]	$3,170	-$4,032		
Subgroup size (P_j)	0.533	0.467		2010-2013 Marketscan[22]
Traditional population ICER (eq. 3)			$46,000/QALY	
Population NMB per patient from statin monotherapy[b]			$315,869	
Adoption of Statin + Fibrate under status quo (f_k = 0.80), D_j	0.072	0.043		2010-2013 Marketscan
Modified population ICER (f_k = 0.80, eq. 7)			$41,733/QALY	
Status-quo policy NMB per patient (f_k = 0.80, eq. 8)			$315,910	
Adoption of Statin + Fibrate under Policy 2 (f_k = 1.0), D_j[c]	0.075	0.045		Estimated[c]
Modified population ICER under Policy 2 (f_k = 1.0, eq. 7)			$41,766/QALY	
Policy 2 NMB per patient (f_k = 1.0, eq. 8)			$315,911	
Adoption of Statin + Fibrate under Hypothetical Policy,[d] D_j	0.23	0.023		Assumed
Population ICER under Hypothetical Policy[d] (eq. 7)			$34,848/QALY	
Hypothetical Policy[d] NMB per patient (eq. 8)			$316,214	

ICER, incremental cost-effectiveness ratios; INMB, incremental net monetary benefit; QALY, quality of adjusted life years.
[a]NMB evaluated at $45,000/QALY for purpose of illustration.
[b]Since at a threshold of $45000/QALY, traditional ICER implies statin + fibrate is not cost-effective and so this is not adopted.
[c]Assuming price-elasticity of demand to be -0.26.
[d]Hypothetical policy assumes f_k = 1.00 and implementation action will increase adoption of the combination therapy by 200% among the males and decrease adoption of combination therapy by 50% among the females from the adoption rate under the status quo policy.

Table 2 Economic Returns from Policy Alternatives based Predicted Diffusion

	Parameters	Policy 1: Status-quo Policy (f_k = 0.80)	Policy 2: Full Coverage (f_k = 1.00)	Policy 3: Full Coverage + Education Program[b]	EDCI eq. (10)	EDPI eq. (11)
Male	P_M	0.533	0.533	0.533	0.533	0.533
	D_M	0.072	0.075	0.23	1	-
	$F(INMB_{AB})$[a]	-	-	-	-	0.32
Female	P_F	0.467	0.467	0.467	0.467	0.467
	D_F	0.043	0.045	0.023	0	-
	$F(INMB_{AB})$[a]	-	-	-	-	0.73
$NMB_{POLICY\ K}$		$315,910	$315,911	$316,214	$317,559	$366,770
LED_{EVPI} (Eq. 16)[c]		-	-	-	$49,211	-
$LED_{POLICY\ K}$		$1,649	$1,648	$1,345	-	-
$LED^*_{POLICY\ K}$		$50,860	$50,859	$50,556	-	-

EDCI, efficient diffusion under current information; EDPI, efficient diffusion under perfect information; NMB, net monetary benefit; ICER, incremental cost-effectiveness ratio; INMB, incremental net monetary benefit; LED, loss in relative to efficient diffusion; LED*, expected incremental loss under a specific policy *k* relative to the Expected Value of Perfection (EVP).

[a]AB represent a comparison between treatment alternatives: here, the combination therapy of statin + fibrate v. the statin monotherapy (i.e., INMB of the combination therapy in relative to the statin monotherapy).

[b]Hypothetical policy assumes f_k = 1.00 and implementation action will increase adoption of the combination therapy by 200% among males and decrease adoption of combination therapy by 50% among females.

[c]LED_{EDPI} is equivalent to the traditional expected value of perfect information (EVPI), representing the opportunity costs of the suboptimal decision caused by existing uncertainty. Also, this can be understood as a maximal value of investing in further research to reduce uncertainty.

Efficient Diffusion under Current Information v. Perfect Information and LED Metrics

Table 2 reports the NMBs under efficient diffusion and current and perfect information along with the policy-specific LED metrics. Under the current information, efficient diffusion would entail the full adoption of the combination therapy among males (i.e., $D_M = 1$ and $D_F = 0$) and no use among females. If achieved, the NMB per patient in the population is estimated to be $317,559 ($NMB_{EDCI}$, Table 2, eq. (10)). However, in this case, we acknowledge the possibility of the suboptimal decision because of potential uncertainty around the expected NMB. Based on the distributions assumed in this stylized example, we found that there was 32% chance of making a wrong decision for males (i.e., 32% chance that the statin therapy would provide better value for males instead of the combination therapy). For females, there was a 27% probability that the combination therapy would rather provide more value than the statin therapy, which is determined by an optimal choice strictly based on the expected NMB threshold. If all uncertainty were to be resolved through future research, we could eliminate the possibility of making a suboptimal decision. Under this scenario of perfect information and perfect implementation, following eq. (11), NMB_{EDPI} is estimated to be $366,770.

The difference between NMB_{EDPI} and NMB_{EDCI}, which represents LED_{EVPI}, is estimated to be $49,211 per patient and reflects the expected value of perfect information with each of the subgroups. The large value of information is driven by the uncertainty in costs that was assumed. Comparing the NMB_{EDCI} to each of the NMB_{Policy} provides a policy-specific estimate of LED, which could be interpreted as the expected value of implementation. Naturally, the policy with the highest NMB_{Policy} also has the lowest LED_{Policy}. To what extent this value represents the value of future implementation v. research would depend on understanding the decision-making rationale, as explained in Section 4 and Appendix A2. The total value of perfection (i.e., perfect implementation and perfect information), under the hypothetical policy, is estimated to be $50,556.

DISCUSSION

We propose a new economic evaluation framework to estimate the realized value of medical technology through a policy lens. First, we argue that standard cost-effectiveness metrics (e.g., ICERs) may be misleading if there are differential adoption rates in different subgroups due to heterogeneity in relative costs or effectiveness. Some have argued that subgroup analyses should be carried out in

CEA when the goal is to provide differential coverage across these subgroups.[13,14] One could conceivably implement subgroup-specific coverage decisions to overcome the limitations of decision making using a single overall ICER. However, in many cases, the number of subgroups could be large, and there may not be easy ways to implement such heterogeneous policies due to ethical and equity concerns and also transaction costs.[10]

Although budget impact analyses (BIA)—as distinct from CEA—have sometimes incorporated potential adoption rates over time to inform affordability, the corresponding impacts on the effectiveness side have largely been ignored.[26] Moreover, even BIA does not typically account for evidence-driven differential adoption across subgroups. Our framework provides summary measures of cost-effectiveness at the population-level, after incorporating subgroup-specific adoption patterns to help population-level decision making when subgroup-specific decision making is not possible.

We highlight the importance of incorporating policy-specific adoption rates of technology into the existing metrics. This framework can be readily applied to compare the economic value of policy options that may often encompass not only binary coverage decisions but also cost-sharing, implementation policies, and investments in research.[27] For example, this framework may provide incentives for manufacturers to provide reliable evidence about heterogeneity so that a single overall ICER based on uniform adoption does not limit the market access for a product that may provide a good value for certain subgroups where differential adoption is anticipated.

We contribute to the literature on heterogeneity in economic evaluation and the value of information approaches. While the expected value of individualized care focused on estimating the value of generating evidence at the individual, decision-making level,[9] our framework emphasizes the role of heterogeneity in population-level (centralized) policy decision making through creating policy-specific metrics. Moreover, we extended the existing literature by introducing policy-specific metrics, rather than treatment-specific metrics, by incorporating adoption behavior induced by such policies, and by allowing for policies (e.g., research investments) to change evidence itself.

There are some limitations to this new framework. First, this framework requires rich information of heterogeneous effects and subgroup-specific results to predict behavioral changes in those groups. Such information is often not readily available. However, with ongoing efforts to increase clinical studies to focus on heterogeneity in comparative effectiveness and patient-centered outcomes, we expect that individual-level patient data and pre-specified subgroup results would provide more useful information for this framework in the near future. We encourage CEA to report subgroup-level results that would help to estimate the realized value of policy decisions more precisely.

Another important limitation is the inability to predict future rates of adoption of new technology. Although we estimated adoption patterns retrospectively, an understanding of the potential changes in adoption in the future would be important to implement this framework for prospective evaluation (e.g., how the adoption of new technology would be changed in response to new clinical evidence). Moreover, we used a constant adoption rate; realistically, one should view the rate of adoption to be time-varying.

Certain policy implementation can also affect supply side factors that can change the adoption of certain technologies. This type of general equilibrium argument is too broad and out of the scope of this paper, but one can evaluate a policy alternative that specifically addresses the supply-side effect and compare policy alternatives based on the demand side.

Finally, in the empirical illustration, we ignore costs of implementing a new policy. However, as with value-of-information methods, one could recognize that implementing a new policy would be worthwhile if the potential gain in the LED metric is greater than the costs of implementation. Also, we presented metrics per patient. In the real application of these methods, results should be presented from a population perspective and for real-policy options, rather than the hypothetical policy alternatives considered in our example.

The above limitations notwithstanding, this novel framework will be useful to evaluate the realized value of medical technology under various decision options. This framework produces metrics that are better aligned with specific policy decisions under considerations for a specific technology, and so can help improve future decisions.

ACKNOWLEDGMENTS

We gratefully acknowledge very the helpful suggestions we received from the editor and three anonymous

reviewers and editorial comments on earlier drafts of this article.

REFERENCES

1. Weinstein MC, Stason WB. Foundations of cost-effectiveness analysis for health and medical practices. N Engl J Med. 1977 Mar 31;296:716–21.

2. Neumann PJ, Sanders GD, Russell LB, Siegel JE, Ganiats TG. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 2016.

3. Neumann PJ, Weinstein MC. The diffusion of new technology: costs and benefits to health care. The Changing Economics of Medical Technology. Washington DC: Institute of Medicine, National Academy of Science; 1991. p 21–34.

4. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold - What it is and what that means. Pharmacoeconomics. 2008;26:733–44.

5. Basu A. Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. J Health Econ. 2011;30:549–59.

6. Basu A, Jena AB, Goldman DP, Philipson TJ, Dubois R. Heterogeneity in action: The role of passive personalization in comparative effectiveness research. Health Econ. 2014;23:359–73.

7. Phelps CE. Good Technologies Gone Bad How and Why the Cost-Effectiveness of a Medical Intervention Changes for Different Populations. Med Decis Making. 1997;17:107–17.

8. Coyle D, Buxton MJ, O'Brien BJ. Stratified cost-effectiveness analysis: a framework for establishing efficient limited use criteria. Health Econ. 2003;12:421–7.

9. Basu A, Meltzer D. Value of information on preference heterogeneity and individualized care. Med Decis Making. 2007;27:112–27.

10. Espinoza MA, Manca A, Claxton K, Sculpher MJ. The value of heterogeneity for cost-effectiveness subgroup analysis: Conceptual framework and application. Med Decis Making. 2014;34:951–64.

11. Dubois RW. From methods to policy: a''one-size-fits-all''policy ignores patient heterogeneity. J Comp Eff Res. 2012;1:119–20.

12. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. Milbank Q. 2004;82:661–87.

13. Meltzer D. Addressing uncertainty in medical cost–effectiveness analysis. J Health Econ. 2001;20:109–29.

14. Sculpher M. Subgroups and heterogeneity in cost-effectiveness analysis. Pharmacoeconomics. 2008;26:799–806.

15. Espinoza MA, Sculpher M, Basu A, Manca A. Analysing heterogeneity to support decision making. In: Culyer T, ed. Encyclopedia of Health Economics; 2014. p 71–6.

16. Fenwick E, Claxton K, Sculpher M. The value of implementation and the value of information: Combined and uneven development. Med Decis Making. 2008;28:21–32.

17. Chernew ME, Rosen AB, Fendrick AM. Value-based insurance design. Health Aff (Millwood). 2007;26:w195–203.

18. Pauly MV, Blavin FE. Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance. J Health Econ. 2008;27:1407–17.

19. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics. 2006;24:1055–68.

20. ACCORD Study Group, Ginsberg HN, Elam MB, et al. Effects of combination lipid therapy in type 2 diabetes mellitus. N Engl J Med. 2010;362:1563–74.

21. Sorensen SV, Frick KD, Wade A, Simko R, Burge R. Model-based simulation to explore the cost-effectiveness of following practice guidelines for triglyceride and low-density lipoprotein cholesterol control among patients with diabetes mellitus and mixed dyslipidemia. Clin Ther. 2009;31:862–79.

22. Danielson E. Health Research Data for the Real World: The MarketScan Databases (White Paper). Ann Arbor, MI: Truven Health Analytics; 2014.

23. Yeung K. Does Cost-Effectiveness Analysis Have a Role in US Managed Care Drug Formularies? An Empirical Study of Utilization, Costs, Outcomes, and Elasticity of Demand of a Value-Based Formulary. Sullivan SD, ed. Seattle, WA: Pharmaceutical Sciences, ResearchWorks; 2015. p 1–62.

24. Leibowitz A, Manning WG, Newhouse JP. The demand for prescription drugs as a function of cost-sharing. Soc Sci Med. 1985;21:1063–9.

25. Ringel JS, Hosek SD, Vollaard BA, Mahnovski S. The Elasticity of Demand for Health Care. A Review of the Literature and its Application to the Military Health System. Rand Monograph Reports. Santa Monica, CA: Rand Corporation; 2002.

26. Sullivan SD, Mauskopf JA, Augustovski F, Jaime Caro J, Lee KM, Minchin M, et al. Budget impact analysis—Principles of good practice: Report of the ISPOR 2012 Budget Impact Analysis Good Practice II Task Force. Value Health. 2014;17:5–14.

27. Griffin SC, Claxton KP, Palmer SJ, Sculpher MJ. Dangerous omissions: the consequences of ignoring decision uncertainty. Health Econ. 2011;20:212–24.