UNDERSTANDING AND ANTICIPATING LAG-TIME BIAS IN COST-EFFECTIVENESS STUDIES: THE ROLE OF TIME IN COST-EFFECTIVENESS ANALYSIS

Gijs van de Wetering
Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands

Marcel Olde Rikkert
Department of Geriatrics, Radboud University Medical Center, Nijmegen, The Netherlands

Gert Jan van der Wilt, Eddy Adang
Department for Health Evidence, Radboud University Medical Center, Nijmegen, The Netherlands
eddy.adang@radboudumc.nl

Background: Timely provision of information on the cost-effectiveness of innovations in health care becomes more and more important, resulting in increasing pressure on researchers to provide proof of cost-effectiveness in a short time frame. However, most of these innovations require considerable time and effort to optimally implement leading to a biased "steady state" cost-effectiveness outcome. As decision makers in health care predominantly have a short-term focus, the discrepancy between short-term study outcomes and long-term cost-effectiveness may very well lead to misguided decisions about the adoption of innovations in health care.
Methods: Factors such as learning effects, capacity constraints, and delayed time to benefit are all related to a short-run timeframe and result in inefficiencies during the implementation of an innovation. These factors and the mechanisms by which they influence the cost-effectiveness outcome are explained for three different types of healthcare innovations.
Results: As standard cost-effectiveness analysis assumes costs and effects to behave constant and representative for an innovation's entire economic lifetime, resulting cost-effectiveness outcomes might give a biased, and often overly pessimistic, reflection of the actual cost-effectiveness of an innovation. This is further amplified by the fact that short-run inefficiencies are most prevalent and impactful during an innovation's earliest stage of operation.
Conclusions: This study advocates to carefully take into account the different factors contributing to lag-time bias in the design and analysis of cost-effectiveness studies, and to communicate potential biases due to short-run inefficiencies to all stakeholders involved in the decision making process.

Keywords: Cost-effectiveness, Time, Inefficiencies, Bias

Nowadays, regulators, funding bodies, and other financing parties are exercising pressure on researchers to provide proof on the cost-effectiveness of an innovative intervention in a (too) short time frame. The question is to what extent outcomes obtained in such a short time frame do justice to the actual cost-effectiveness of a new intervention. The majority of innovations in health care are complex in nature, meaning that they require considerable time and effort to implement ([1];[2]). Even for rather common drugs such as statins, and (novel) anticoagulant agents it is clear that "time to benefit" is a crucial, although highly neglected characteristic in medical decision making ([3]). Hence, there might be reason to believe that a discrepancy exists between short- and long-run cost-effectiveness outcomes ([4]). As the outcomes of decision analytical modeling are largely dependent upon clinical trials that use a short time frame for their evidence generation, researchers should make a strong effort to provide insight into the usability of these short-run outcomes for funders and decision makers when determining the cost-effectiveness of an innovation. The aim of this paper is to provide insight in the mechanisms by which short-run inefficiencies arise and how a (too) short study duration in-

fluences the reliability of cost-effectiveness (CE) outcome in studies of new health technologies.

Short-Run Inefficiencies in Cost-Effectiveness Analysis

Cost-effectiveness analysis (CEA) as part of the evaluation of medical innovations, both in drug and nondrug interventions, has become well accepted and widely applied in several high-income countries. For example, in the United Kingdom, the National Institute of Clinical Excellence (NICE) uses cost-effectiveness outcomes as a criterion for coverage recommendations to the National Health Service. Theoretically standard CEA should be performed in a common practice setting and aims to collect all relevant costs and effects for all alternative interventions (mutually exclusive) over a representative time period. However evidence on cost-effectiveness is often gathered alongside clinical trials or is derived from available evidence by decision analytical modeling. Cost-effectiveness of a new intervention relative to current practice is often expressed in a decision rule relating the incremental cost-effectiveness ratio (ICER) to some societal willingness to pay for a unit of effect like for example a quality-adjusted life

year gained. In using this ICER, costs and effects collected during the trial period are averaged and assumed to be constant over and representative for the economic lifetime of the intervention. However, this assumption is not always realistic as it neglects possible inefficiencies that occur during the implementation phase, where additional costs are often unanticipated and clinical effectiveness for the innovation is often not yet optimal (5). As most innovations are in a transitional state when evaluated, inefficiencies are likely to occur, and the resulting ICER might be an incorrect reflection of the actual cost-effectiveness of the innovation. The impact of these inefficiencies on the outcomes of economic evaluations in health care is strongly related to the timeframe available for evaluation. It is generally assumed that if the timeframe is sufficiently long than the CEA outcome is a reliable reflection of the "steady state". However, as argued before this is often not the case. In order to be able to estimate the impact of short-run inefficiencies, and thus assess the expected accuracy of a CEA, there needs to be transparency about the underlying factors that might cause a discrepancy between the estimated ICER and the "true" steady state ICER. Here, three major drivers of short-run inefficiencies are identified and discussed: learning effects, capacity constraints, and delayed time to benefit. The choice for these three factors is based on micro economic theory (short-run versus long-run cost behavior) and research methodological reasons (a timeframe representative for the long-term CE outcome). Also, when one or more of these factors are present, the consequences for the outcomes of an economic evaluation often become clinically relevant, as the actual decision whether or not to adopt an innovation is easily changed by the presence of these factors (5).

Learning Effects

Learning can both take place at an individual as well as at an organizational level (6). Learning effects are particularly visible in interventions where human actions have significant influence on procedural outcomes, for example when performing innovative surgical procedures or a switch from a 2D to a new 3D diagnostic modality. Learning effects arise because initially clinical personnel involved in a new and complex procedure will perform suboptimal. This suboptimal performance will manifest itself in longer procedure times, higher complication rates, and lower quality of life outcomes and survival rates for patients, as compared to the performance of the innovation when the learning curve has reached a plateau (7). As the usual care is usually free of any learning effects, not correcting for learning can lead to an inaccurate estimation of the incremental cost-effectiveness of the innovation. A well-known example is laparoscopic surgery, in which learning tends to exert a large influence on the results of clinical trials (8). For example, Broeders et al. studied the impact of surgeon experience on 5-year outcome of laparoscopic Nissen fundoplication (LNF) for gastroesophageal reflux disease (9). This study found a significant decrease in operating time, in-hospital complications, early

dysphagia, and conversions from laparoscopic to conventional surgery. Here the impact of learning on trial outcomes and its consequences can be directly related to the original randomized controlled trial (RCT) which was prematurely terminated because seven patients in the LNF group had developed dysphagia versus none in the control group (10). The study of Broeders et al. confirmed the criticism on the RCT that not taking into account learning had introduced bias causing inferior outcomes for the laparoscopic treatment arm. Researchers should be careful when performing a trial in a multicentre setting consisting of centers of excellence and less experienced centers. Depending on the distribution of the inclusion of patients between these centers, as well as the absolute number of patients treated in these centers, the resulting ICER might not be representative for the steady state ICER.

Capacity Constraints

When substituting one intervention for another, it is often assumed that the old intervention can be scaled down at the decision maker's own discretion or be fully replaced without suffering financial losses due to excess capacity (11). This is, however, dependent on the rate of implementation and whether the technologies under evaluation are convex (perfectly divisible). When an innovation is rapidly implemented, not all resources of the old intervention can immediately be freed up to finance the innovation. For example, medical equipment not yet at the end of their economic life (e.g., older generation CT scanners) decrease the anticipated savings in the short run (12). This in turn increases the ICER in the short run. On the other hand, when implementation commences in a more gradual manner, for example due to organizational and infrastructural changes that need to be made, the innovation may not have enough capacity to supply all patients at once. This results in the need to keep the old and relatively inefficient intervention operational throughout the implementation period, resulting in both diseconomies of scale and the healthcare provider not reaping the full benefits of the new and more cost-effective intervention. An illustrative example is the conversion from analogue screen film mammography to digital mammography, which is a process that takes place in a gradual manner as the technological substitution takes place on a nationwide level (13).

Delayed Time to Benefit

Even when implementation of an innovation is optimal, a clinically meaningful therapeutic effect is not guaranteed to be immediately present. An example is statins, for which "time to benefit" has emerged as a useful concept to assess efficacy in specific settings over varying periods of time (14). For example, treatments that lower LDL cholesterol in patients with coronary artery disease reduce these LDL levels dramatically within months. However, an actual reduction in angiographic disease progression may require 3 years (15;16). This has obvious implications for cost-effectiveness analysis where the main focus lies on final outcome measures such as quality of life

Table 1. Influence of Factors Driving Short-Run Inefficiencies on Different Types of Healthcare Interventions

	Pharmaceutical	Surgical	Diagnostic
Learning effects	++	++	+
Capacity constraints	−−	+	++
Delayed time to benefit	++	+	−−

Note. ++ influence can be very strong and/or is prevalent in intervention type; + influence can be strong but is relatively rare in intervention type; −− influence is negligible and/or very rare/non-existent in intervention type.

and mortality and where the exact functional behavior between the intermediate (LDL level) and final outcome (quality of life, survival) are not precisely known. For innovations with long time to benefit on these outcome measures, cost-effectiveness may be underestimated during the evaluation period when the focus is on the final outcomes or might be imprecise because of misspecification of the relationship between intermediate and final outcome. This particular problem of time to benefit might be alleviated by correcting for the fact that a trial only gives intermediate endpoints. By correctly extrapolating intermediate endpoints in a trial to final endpoints to be used in an economic evaluation, research conclusions can be nuanced while still serving as a solid guide for decision making about policy and further research. However, often clinical trials are not designed to provide input for decision analytical models. Therefore, researchers should pay specific attention to the usability of trial data in their models, especially with regard to the relationship between intermediate trial outcomes and model input parameters, specifically the outcomes over time. The inclusion of intermediate outcomes in health economic models should acknowledge the uncertainty surrounding this relationship. Hence, this problem should be tacked using simple methodology, such as including the model parameters contingent on intermediate outcome measures in deterministic sensitivity analyses.

Besides the three drivers presented here, there are other factors that can trigger short-run inefficiencies. However, most of these factors can be traced back to the three drivers presented in this study. When looking at the nature of healthcare innovations we distinguish three main groups in this article, namely pharmaceuticals, surgical interventions, and diagnostic interventions. Table 1 provides an overview of these innovations and a description of the extent in which learning effects, capacity constraints, and delayed time to benefit drive short-run inefficiencies for these particular innovations, based on analysis of the examples given.

When comparing these three types of interventions, We see that learning might play a role for all interventions. For pharmaceuticals, learning might both play a role at the introduction (when professionals only slowly learn the rare side effects, that cannot be found in smaller RCTs) and also later on when

dosages and packing are adjusted. Unfortunately in several drug developments of recent years (e.g., Cox-2 inhibitors, new oral antidiabetics, the dpp-4 inhibitors), the cost-effectiveness estimation changed considerably from introduction on the market to the appearance of much less positive phase IV surveillance data (17;18). Regarding capacity constraints, these do not play a role for pharmaceuticals, as their cost behavior is mostly variable, but are a concern when adopting innovations in the surgical or diagnostic area. Delayed time to benefit can obviously impact pharmaceuticals, but also plays a role in surgical interventions, as described by Buchwald et al. (15;16).

Ideally, all these factors should be modeled in one overarching algorithm that gives, depending on the nature of the innovation, a correction to the cost-effectiveness outcome. Due to the variety between innovations, this might be rather difficult. However, existing outcomes research might be improved to better capture foreseen short-run inefficiencies. For now, the most important task for health economic researchers is to raise awareness between themselves and decision makers about the possible implications of these short-run inefficiencies.

Consequences of Biased Cost-Effectiveness Outcomes

The failure to address the short-run–long-run discrepancy in standard CEA can have important consequences. Most decision makers and funders tend to focus on short-term results and demand these results on short notice (19). This myopic attitude reduces the length of trial periods and outcome measurements, as researchers feel pressured to shorten the time frame for their studies. Additionally, researchers may display myopic behavior themselves. An example is when health economic models are populated with data that are generated within a too short timeframe with regard to cost-effectiveness, and are hence not representative for the steady state health economic outcome. As short-run inefficiencies are most prevalent and impactful during an innovation's earliest stage of operation, reducing the length of evaluation periods will result in cost-effectiveness outcomes being biased even more against the innovation. If no attempt is made to correct for this by using for example intermediate outcome measures or taking potential short-run inefficiencies into account in the interpretation of study results, this might lead to outcomes that seriously underestimate the cost-effectiveness of a new intervention, potentially resulting in second thoughts about the implementation of an in essence cost-effective innovation. Worst-case scenario is that such a "lag-time bias" in cost-effectiveness analysis will result in denying patients access to more efficient health care. Of course the opposite may also be true, in the sense that cost-effectiveness will decrease later on, for example by extending the indication to other patient groups with less benefit. This is for example suspected in the recent case of transcatheter aortic valve replacement, but as far as we know, it has not been demonstrated in cost-effectiveness data (20).

Careful health technology assessment (HTA) of new medical interventions is increasingly important and more and more

610

considered as obligatory for further implementation and re-imbursement of healthcare innovations. Coming back to the research question, we argue that the combination of a myopic focus of decision makers and the discrepancy between study outcomes and long-term (steady state) cost-effectiveness of an innovation can lead to misguided decisions about the implementation of innovations in health care, due to short-run inefficiencies surrounding the substitution of the old technology. It should be acknowledged that if the dynamics in a certain health market are large and innovations follow in rapid succession a long-run steady state of a particular innovation will probably not be reached. However, such a market is unlikely to exist for a long period as, in the context of large discrepancies in efficiency between the short run and long run, healthcare providers will strive to produce their products on the long-run efficient frontier because this gives them a competitive advantage over those that do not exploit economies of learning and scale. The mechanism by which short-run inefficiencies manifest themselves is dependent on a multitude of factors, most importantly the type of innovation and the characteristics of the to-be-substituted technology. The role of short-run inefficiencies in the evaluation of an innovation offers ample opportunities for further research. We, therefore, urge policy makers and researchers to expand research in this important, but relatively neglected topic. HTA studies that extrapolate short-term outcomes may fail to deliver good estimates of cost-effectiveness, often underestimating the real effects of an innovation. Therefore, we advocate to carefully take into account the different factors that may cause this lag-time bias in the design and analysis of cost-effectiveness studies. Here lies a crucial responsibility for researchers in the field of health economics, as they are the ones being able to raise awareness among decision makers on these methodological issues. Therefore, it is essential that information on potential short-run inefficiencies is reported alongside the results of an economic evaluation. If this information is communicated to decision makers, together with options to deal with these inefficiencies, it can help to prevent pitfalls in implementation of innovations, and guide important decisions on the implementation strategy.

CONTACT INFORMATION

Gijs van de Wetering, PhD, Department for Health Evidence, Radboud University Medical Center, Nijmegen 6500 HB, The Netherlands

Marcel Olde Rikkert, Prof.dr, Department of Geriatrics, Radboud University Medical Center, Nijmegen 6500 HB, The Netherlands

Gert Jan van der Wilt, Prof.dr, Eddy Adang, PhD (eddy.adang@radboudumc.nl), Department for Health Evidence, Radboud University Medical Center, Nijmegen 6500 HB, The Netherlands

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Reelick MF, Faes MC, Esselink RA, Kessels RP, Olde Rikkert MG. How to perform a preplanned process evaluation for complex interventions in geriatric medicine: Exemplified with the process evaluation of a complex falls-prevention program for community-dwelling frail older fallers. *J Am Med Dir Assoc*. 2011;12:331-336.
2. Faes MC, Reelick MF, Esselink RA, Rikkert MG. Developing and evaluating complex healthcare interventions in geriatrics: The use of the medical research council framework exemplified on a complex fall prevention intervention. *J Am Geriatr Soc*. 2010;58:2212-2221.
3. Holmes HM, Hayley DC, Alexander GC, Sachs GA. Reconsidering medication appropriateness for patients late in life. *Arch Intern Med*. 2006;166:605-609.
4. Adang EM, Voordijk L, Van der Wilt GJ, Ament A. Cost-effectiveness analysis in relation to budgetary constraints and reallocative restrictions. *Health Policy*. 2005;74:146-156.
5. van de Wetering G, Woertman WH, Adang EM. Time to incorporate time in cost-effectiveness analysis. *Eur J Health Econ*. 2012;13:223-226.
6. Argote L, Epple D. Learning Curves in Manufacturing. *Science*. 1990;247:920-924.
7. Anzanello MJ, Fogliatto FS. Learning curve models and applications: Literature review and research directions. *Int J Ind Ergon*. 2011;41:573-583.
8. Ramsay CR, Grant AM, Wallace SA, et al. Assessment of the learning curve in health technologies. A systematic review. *Int J Technol Assess Health Care*. 2000;16:1095-1108.
9. Broeders JA, Draaisma WA, Rijnhart-de Jong HG, et al. Impact of surgeon experience on 5-year outcome of laparoscopic Nissen fundoplication. *Arch Surg*. 2011;146:340-346.
10. Bais JE, Bartelsman JF, Bonjer HJ, et al. Laparoscopic or conventional Nissen fundoplication for gastro-oesophageal reflux disease: Randomised clinical trial. *Lancet*. 2000;355:170-174.
11. Smet M. Cost characteristics of hospitals. *Soc Sci Med*. 2002;55:895-906.
12. van de Wetering G, Woertman WH, Adang EM. A model to correct for short run inefficiencies in health care. *Health Econ*. 2012;21:270-281.
13. van de Wetering G, Woertman WH, Verbeek AL, Broeders MJ, Adang EM. Quantifying short run cost-effectiveness during a gradual implementation process. *Eur J Health Econ*. 2013;14:911-918.
14. Ray KK, Cannon CP. Time to benefit. An emerging concept for assessing the efficacy of statin therapy in cardiovascular disease. *Crit Pathw Cardiol*. 2005;4:43-45.
15. Buchwald H, Moore RB, Rucker RD Jr, et al. Clinical angiographic regression of atherosclerosis after partial ileal bypass. *Atherosclerosis*. 1983;46:117-128.
16. Buchwald H, Varco RL, Matts JP, et al. Effect of partial ileal bypass surgery on mortality and morbidity from coronary heart disease in patients with hypercholesterolemia. Report of the Program on the Surgical Control of the Hyperlipidemias (POSCH). *N Engl J Med*. 1990;323:946-955.
17. Chancellor JV, Hunsche E, de Cruz E, Sarasin FP. Economic evaluation of celecoxib, a new cyclo-oxygenase 2 specific inhibitor, in Switzerland. *Pharmacoeconomics*. 2001;19(Suppl 1):59-75.
18. Brown TJ, Hooper L, Elliott RA, et al. A comparison of the cost-effectiveness of five strategies for the prevention of non-steroidal anti-inflammatory drug-induced gastrointestinal toxicity: A systematic review with economic modelling. *Health Technol Assess*. 2006;10:iii-iv, xi-xiii, 1-183.
19. Drummond M, Cooke J, Walley T. Economic evaluation under managed competition: Evidence from the UK. *Soc Sci Med*. 1997;45:583-95.
20. Van Brabandt H, Neyt M, Hulstaert F. Transcatheter aortic valve implantation (TAVI): Risky and costly. *BMJ*. 2012;345:e4710.