Paying for cost-effective health care: Does it violate both static- and dynamic efficiency?

TOMAS J. PHILIPSON [1, *]

[1] University of Chicago, United States

Abstract: Recent research has questioned the rationales of using cost-effectiveness metrics of medical technologies to guide reimbursement. I discuss here the underlying ideas of this research, which argues that reimbursement based on cost effectiveness criteria leads to both static- and dynamic inefficiencies.

JEL Classification: I1, I11, I13

Key words: pharmaceuticals, reimbursement, cost-effectiveness analysis

1 Introduction

Growth in health care spending across the world has received much attention, particularly due to its impact on public budgets and national debt levels. In assessing the causes of this growth, new medical technologies are often argued to be leading forces behind the growth in both private and public health care spending. In order to manage the spending growth induced by new technologies, both public and private payers around the world have increasingly demanded evidence on the combined measures of the benefits and costs of new technologies. These measures include, among others things, cost-effectiveness, cost-utility, and cost-benefit analysis; they are hereafter referred to collectively as CE analysis. Indeed, the amount of work done on the CE of medical technologies may perhaps be the largest field within health economics, particularly in European countries where such analysis already guides a large share of public technology adoption and reimbursement. In practice, CE analysis has so far guided policy decisions in the form of adoption based on CE thresholds, which dictate that a given technology will be reimbursed only if the incremental costs per quality-adjusted life year (QALY) they provide are below a given threshold. As is well known, CE analysis already plays a role in public reimbursement decisions outside the US. For example, both the UK's National Institute for Clinical Excellence (NICE) and Australia's Pharmaceutical Benefits Advisory Committee have explicitly invoked CE thresholds in technology adoption decisions. The Nordic countries do not rely exclusively on formal CE thresholds, e.g. the Swedish case TLV considers uncertainty in data and severity of illness and therefore reimbursement decisions show a wide range of cost per QALY.

Although it is well known that CE analysis is less institutionalized in the more privately financed US market, the majority of worldwide CE studies are done for the US market and funded by US manufacturers. The interest by manufactures in funding such studies suggests that payers are influenced by CE studies even in the US market place. The core and seemingly self-evident assumption behind using CE based analysis is that payers

Correspondence to: Tomas J. Philipson, Becker Friedman Institute and Irving B. Harris Graduate School of Public Policy Studies, The University of Chicago, 1155 E. 60th Street, Chicago IL 60637. E-mail: t-philipson@uchicago.edu

dx.doi.org/10.5617/njhe.1289

should attempt to maximize the returns in health they obtain from the limited resources available for health spending. The methods of CE analysis are the main tools proposed to guide payers to do this.

While CE analysis has provided a guide to allocating often-scarce health care resources, less is known about the exact degree to which using such procedures results in economically efficient allocations. Recent literature has raised two major concerns regarding the efficiency effects of evaluating health care interventions through CE analysis. First, when utilized in practice by payers, do they lead to the resource allocation decisions that satisfy static efficiency as intended? Second, if CE procedures guided reimbursement as prescribed, would they lead to resource allocations that satisfied dynamic efficiency? Recent research suggests that the answers may be "no" to both questions.

2 Static efficiency and endogenous cost-effectiveness analysis

Static efficiency results if other things constant, e.g. severity of illness and size of the patient population, the treatments that are cheaper to produce are utilized more in health care delivery. Jena and Philipson (2013) argue that the standard static efficiency rationale for using CE criteria to adopt technologies breaks down when producers act optimally, faced with these adoption procedures. This occurs because CE analysis uses the prices charged to payers by producers, rather than the resource costs used in production, which drives static efficiency. This is an inevitable part of technology adoption, as producers in any industry are reluctant to share data on production costs. Therefore, endogenous prices determine the CE levels observed for new innovations, not the production costs that ordinarily determine the efficient use of resources. The key implication of this is that observed CE levels, being determined by prices, are the result of how CE analysis is used by payers in setting reimbursement. For example, if a payer only pays for technologies that are cheaper than a fixed CE threshold, as e.g. often argued to be the goal in the UK, manufacturers may find it in their best interest to set prices up to that threshold regardless of production costs. Thus, treatments with different production costs may appear equally cost-effective due to the particular CE threshold adoption policy. Treatments appear equally cost-effective despite the great variation in the resources used to produce them. As they are equally cost-effective, there is no selection of good technologies over bad ones because CE levels are equalized due to the optimal pricing given the CE-based reimbursement rule used. This problem occurs even if other dimensions than QALYs are used to determine what acceptable prices are. For example, if, as in Sweden, uncertainty in data and severity of illness are other attributes determining acceptable prices, manufacturers will price as high as the QALYS gained and the consideration of those factors allows for. Holding constant the multiple factors, treatments with greatly varying production costs will be priced and adopted in a similar manner.

Thus, utilization of cost-effectiveness- or other types of analysis for reimbursement purposes is subject to a form of the "Lucas critique"; the stated goals of the policy will not materialize when those affected by the policy act optimally when faced with it. Put simply, the standard rationale of picking technologies based on CE levels within a given budget breaks down when CE levels are endogenous to how the technologies are picked.

This has two important implications for assessing the impact of CE analysis in guiding health care resource decisions. First, under endogenous cost-effectiveness levels induced by optimal pricing, policies aimed at raising cost-effectiveness may actually lower it. This is because pricing may respond in unanticipated ways to the more stringent adoption procedures. Second, reimbursement policy based on endogenous cost-effectiveness levels may lead to the adoption of more inefficient treatments. The overall argument is that the

static efficiency rationale for using CE assessments for health care adoption is weakened when those affected by such adoption policies act in their own self-interest.

3 Dynamic efficiency and reimbursement based on cost-effectiveness analysis

Dynamically efficient pricing of treatment occurs when all costs and benefits over time, not only current ones, are taken into account. Little analysis exists on the effects of using CE analysis for technology adoption on dynamic efficiency. However, a better understanding of the link between innovation and cost-effectiveness analysis is particularly important given the large role of technological change in the growth of health care spending as well as the growing use of CE thresholds in guiding technology adoption in several countries. Indeed, while cost-effectiveness (CE) analysis has provided a guide to the static allocation of scarce health care resources, less emphasis has been placed on its effect on the behavior of innovators who make health care technologies available in the first place.

Jena and Philipson (2008) stresses that an important aspect of CE-based technology adoption is that it closely resembles other forms of supply price regulations, such as price controls and rate-of-return regulations, and therefore has similar implications for dynamic efficiency. CE thresholds, which are termed "buyer reservation prices" in standard economic language, are price controls in the sense that if price determines costs and the health-effects of products determine effectiveness, adoption policies based on cost-effectiveness are adoption policies based on publicly controlled prices. Although not explicitly stated as such, CE thresholds utilized in practice are implicitly concerned with maximizing the surplus available to consumers at the cost of reduced producer surplus or profits to innovators. In particular, various forms of CE assessments attempt to quantify the health impacts of new technologies for patients by comparing patient benefits from a given technology with total spending on that technology. The central theme of such standard CE assessments performed in practice seems to be to measure consumer surplus or net consumer benefits—technologies are deemed more valuable the larger the patient health benefits, however measured, relative to what is spent on them.

However, when new technologies are brought to life from costly R&D, consumer surplus may be a poor guide to inducing optimal R&D investments, and maximizing static efficiency may harm dynamic efficiency. Rather, the degree to which producer surplus captures social surplus, often at the expense of consumer surplus, becomes the central issue that determines dynamic efficiency. This, of course, is the rationale for the patent system, which replaces consumer surplus with producer surplus in order to stimulate dynamically efficient R&D investment. For the same reason that patents are preferred, even though they lower consumer surplus and CE after technologies are discovered, technology adoption criteria are preferred that do not just focus on consumer surplus as CE criteria do. Put differently, even though measured levels of CE would be larger without patents, since patients or health plans would spend less to get the same technology everyone agrees this would not be desirable as dynamic efficiency would presumably be damaged. An illustrative case of this may be childhood vaccines, which, due to government monopsony power, many times have been estimated to be extremely cost-effective yet lack any appreciable R&D investments as a consequence. The other side of the coin is that many analysts have argued that recent oncology treatments are highly cost-ineffective, though it is perhaps the most active R&D area as a consequence.

To illustrate the dynamic inefficiencies of using CE-based reimbursement, consider the simplest case of patent protected monopoly R&D. In this case it may be that static efficiency and dynamic efficiency, as well as patient health, are maximized when cost-

effectiveness is minimized, that is, when the price per QALY is at its highest feasible level. This occurs under perfect price discrimination. It is well known that in this case, the monopolist gets the entire social surplus, which thus makes his R&D investments dynamically efficient as both the full costs and benefits of those investments are internalized by the monopolist. Patient health is maximized as all consumers buy the product, that is, there is no monopoly distortion of output by restricting it. However, prices are as high as they can be and cost-effectiveness levels are therefore the worst or low as they can be.

The overall point these two major efficiency issues raises is that much more research needs to be done on the implications of using CE-based reimbursement policies for technology adoption. Arguing that cost-effective technologies should be adopted by private- or public payers may lead to both static and dynamic inefficiencies that may harm patients paying in premiums or taxes for their own care.

References

Chernew, M. and Newhouse, J. (2012). Health care spending growth. *Handbook of Health Economics*, Volume 2, edited by Mark V. Pauly, Thomas G McGuire, Pedro Pita Barros, North-Holland.

Drummond, M. F., O'Brien, B., Stoddart, G. L and Torrance, G.W. (1997). *Methods for the economic evaluation of health care programmes.* Oxford: Oxford University Press.

Garber, A. and Pehlps, C. (1997). Economic foundations of cost-effectiveness analysis. *Journal of Health Economics,* 16(1), 1-31.

Gold, M. R., Siegel, J. E., Russell, L. B. and Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine.* Oxford: Oxford University Press.

Jena, A. and Philipson, T. (2008). Cost-effectiveness analysis and innovation. *Journal of Health Economics*. 27(5), 1224-1236.

Jena, A. and Philipson, T. (2013). Endogenous cost-effectiveness analysis and health care technology adoption. *Journal of Health Economics*, 32(1), 172-180.

Johannesson, M. and Weinstein, M. C. (1993). On the decision rules of cost-effectiveness analysis. *Journal of Health Economics*, 12(4), 459-467.

Meltzer, D. (1997). Accounting for future costs in medical cost-effectiveness analysis. *Journal of Health Economics*, 16(1), 33-64.

Weinstein, M. C. and Stason, W. B. (1997). Foundations of cost-effectiveness analysis for health and medical practices. *New England Journal of Medicine*, 296(13), 716-721.