Tulsa Law Review

Volume 49
Issue 3 *Health Law Policy - Honoring the Work of Einer Elhauge*

Article 3

Spring 2014

The Presumption against Expensive Health Care Consumption

Christopher T. Robertson

Follow this and additional works at: http://digitalcommons.law.utulsa.edu/tlr

Part of the Law Commons

Recommended Citation
Christopher T. Robertson, *The Presumption against Expensive Health Care Consumption*, 49 Tulsa L. Rev. 627 (2014).

Available at: http://digitalcommons.law.utulsa.edu/tlr/vol49/iss3/3

This Legal Scholarship Symposia Articles is brought to you for free and open access by TU Law Digital Commons. It has been accepted for inclusion in Tulsa Law Review by an authorized administrator of TU Law Digital Commons. For more information, please contact daniel-bell@utulsa.edu.

THE PRESUMPTION AGAINST EXPENSIVE HEALTH CARE CONSUMPTION

Christopher T. Robertson *

This essay, as part of a symposium in honor of Professor Einer Elhauge, starts with his recognition that, for both epistemic and normative reasons, it remains profoundly difficult to regulate particular uses of medical technologies on the basis of their cost-benefit ratios. Nonetheless, this essay argues in favor of a general regulatory presumption against consumption for the most expensive medical technology usages, which drive most of aggregate healthcare spending. This essay synthesizes twelve facts about the ways in which medical technologies are produced, regulated, studied, and consumed to suggest that it is quite unlikely that the most expensive usages of medical technologies will have benefits exceeding their costs. These considerations include the contingent relationship between research investments and health outcomes, the Food and Drug Administration's ("FDA") lack of authority to consider cost-effectiveness, and the prevalence of off-label uses that have no proof of efficacy. Where efficacy is shown, there are problems in scientific research including publication biases, lack of effective randomization and blinding, commercial biases, the use of surrogates for improved health, small demonstrated benefits not enjoyed by most consumers, and the lack of power and time to detect adverse outcomes. There are also market failures because consumers are unable to estimate benefits and have little or no exposure to cost, while their advisors, physicians, have misaligned incentives. Ultimately, aggregate data across time, geography, and experimental conditions shows that much medical spending is along the "flat of the curve," not delivering commensurate healthcare value. Thus even without particularized rationing decisions, crude regulatory tools that reduce consumption, while preserving choice, are likely to promote rather than hinder welfare.

INTRODUCTION

"[W]e could today easily spend 100% of our GNP on health care without running out of services that would provide some positive health benefit to some patient . . . . [T]radeoffs must be made."[1] In a landmark 1994 article, Einer Elhauge took this problem head on. He rejected simplistic market mechanisms, acknowledging that, "[a]n individu-

* Christopher Robertson, J.D., Ph.D, is an Associate Professor at the James E. Rogers College of Law, University of Arizona, and wrote this piece will visiting at Harvard Law School and affiliated with the Edmond J. Safra Center for Ethics at Harvard University. The author appreciates helpful comments from other attendees at the University of Tulsa College of Law symposium, those from Glenn Cohen and the participants at the Petrie Flom Center Health Law Policy Workshop, and those from Ameet Sarpatwari.

1. Einer Elhauge, *Allocating Health Care Morally*, 82 CAL. L. REV. 1449, 1459 (1994).

Published by TU Law Digital Commons, 2014

al's ability to pay should indeed be irrelevant to determining that individual's access to the minimum of adequate care."[2] Elhauge argued that we must make a political decision about how much of our social resources to allocate towards health care, but within any health plan, there will remain difficult questions of healthcare allocation, ones that "affect profound matters of life, death, and health."[3] Accordingly, Elhauge argues that each individual should be allowed to select from a diversity of health insurance plans, each reflecting whatever rationing priorities a group of individuals may prefer. Those rationing policies would be applied by technocrats: "professionals who have the range of diagnostic expertise to evaluate the healthcare needs of plan enrollees."[4] To avoid gaming of these choices, Elhauge suggests theories of consent that would allow policymakers to bind individuals to the plans they chose *ex ante*.

Twenty years later, the profoundly difficult questions about allocation of healthcare remain. However, the envisioned diversity of health plans has not arisen to allow self-clustering of beneficiaries around chosen rationing rules. Instead, we find ourselves muddling along in large insurance pools, provided by governments or employers to people who have been exogenously clustered by contingencies of geography and firm organization. These public and private insurers have largely punted on the rationing imperative, and now pay for all sorts of high-cost treatments with little or no proven efficacy, and without any serious concern for cost-effectiveness as such.[5] In America, market forces, along with policymakers and courts, still reject the notion that insurers, physicians, or the government should ration on our behalf.[6]

In a 1996 article, Elhauge presciently diagnosed this difficulty of breaking the healthcare system out of this "absolutist imperative that encourages the provision of all medical care having positive net health benefits regardless of cost."[7] Elhauge worried

2. *Id.* at 1455.

3. *Id.* at 1453.

4. *Id.*

5. *See* Christopher T. Robertson, *The Split Benefit: The Painless Way to Put Skin Back in the Health Care Game*, 98 CORNELL L. REV. 921, 935-939 (2013). *See also* James D. Chambers, Peter J. Neumann, & Martin J. Buxton, *Does Medicare Have an Implicit Cost-Effectiveness Threshold?* 30 MEDICAL DECISION MAKING E14 (July-Aug 2010) (concluding that "CMS is covering a number of interventions that do not appear to be cost effective," and do not appear to be using a cost-effectiveness threshold). There are noteworthy anecdotes and exceptions to this rule. *See e.g.*, E. Ray Dorsey & David O. Meltzer, *The Economics of Comparative Effectiveness Research*, 75 NEUROLOGY 492, 493 (2010):

> In the 1990s, the then Agency for Health Care Policy and Research funded a study of the management of back pain and found that much back surgery was unnecessary. The response was not a move to less expensive (and more effective) treatment options but rather intense lobbying by vested interests and ultimately, a move to reduce funding for the Agency.

See also Stephen S. Hall, *The Cost of Living*, N.Y. MAGAZINE, (Oct. 21, 2013), nymag.com/news/features/cancer-drugs-2013-10/#print (describing how the formulary committee for Sloan-Kettering hospital recently decided, for the first time ever, not to provide a drug, because of its price).

6. *See* William M. Sage, *Should the Patient Conquer?*, 45 WAKE FOREST L. REV. 1505, 1510 (2010) ("[W]e regard health insurance as a life raft for those in peril instead of a common-pool resource requiring stewardship. We reach desperately for any new technology that might help defeat death. Any preplanned limit seems like a death panel."); DAVID MECHANIC, FROM ADVOCACY TO ALLOCATION: THE EVOLVING AMERICAN HEALTH CARE SYSTEM 215 (1986) (It is "highly unlikely that the American population would support the rationing of expensive high technology in the fashion characterizing England's National Health Service.").

7. Einer Elhauge, *The Limited Regulatory Potential of Medical Technology Assessment*, 82 VA. L. REV.

that this form of healthcare financing incentivizes the production of "innovations that marginally improve medical outcomes despite great cost."[8] Although policymakers could try to screen these low-value technologies from reaching the market, Elhauge wisely "doubt[ed]" that "regulation can do much about the cost escalation problem. The technology we get reflects the incentive structure for using it."[9] Elhauge explained that the task of regulating particular medical technologies—to distinguish between the worthwhile and the wasteful—foundered on normative and epistemic problems.

Epistemically, regulators often lack the rigorous scientific information that would be necessary to quantify the causal benefits of medical technologies. Elhauge explains that such regulators "face enormous technical problems" as the costs and benefits "vary with output and regions, among individuals and across time. No centralized regulator could possibly implement regulations that effectively adjust for all these factors and shift quickly with time and region as the factors change."[10]

Normatively, even assuming a regulator could reliably identify the marginal benefits of a medical technology in a particular use, she still must assess whether the health benefits are worth the price. That task requires her to put a dollar value on an extra few days of life, or on sight versus blindness. Although routinely done in other countries, these questions require "value judgments that are not susceptible to objective scientific determination."[11]

For example, consider the thirteen new cancer drugs approved in 2012. Twelve of them "were priced above $100,000 annually."[12] When used for the FDA-approved condition, only one of these drugs "provides survival gains that last more than 2 months," when tested in the carefully-constructed clinical trial setting.[13] Scholars have thus estimated that these drugs cost about $33,500 for each additional month lived.[14] Even with these facts stipulated, and the sense of aghast that many infect some readers, it is still difficult to say that these are such bad tradeoffs that regulators should ban them and insurers should refuse to pay for them (binary decisions to allow or disallow the technology).[15] But even worse, when these new drugs are used off-label, as is frequently common in the

1525, 1526 (1996).

8. *Id.*

9. *Id.*

10. *Id.* at 1530. Similarly, see Adam G. Elshaug, J. Michael McWilliams, & Bruce E. Landon, *The Value of Low-Value Lists*, 309 J. AM. MED. ASS'N 775, 775 (2013) (discussing patient heterogeneity as an obstacle to cost-benefit evaluation).

11. Elhauge, *supra* note 7, at 1529.

12. Donald W. Light & Hagop Kantarjian, *Market Spiral Pricing of Cancer Drugs*, 119 CANCER 3900, 3900 (2013).

13. *Id.*

14. Camille Abboud et al., *The Price of Drugs for Chronic Myeloid Leukemia (CML) is a Reflection of the Unsustainable Prices of Cancer Drugs: From the Perspective of a Large Group of CML Experts*, BLOOD (Apr. 25, 2013), http://bloodjournal.hematologylibrary.org/content/early/2013/04/23/blood-2013-03-490003.full.pdf ("[f]or example, in pancreatic cancer, where the median survival is 6 months, a new drug that may prolong survival by 2 months, and is priced at $100,000 per year, will cost $67,000 over 8 months survived, or $33,500 per additional month lived, equivalent to $400,000 per additional year lived.").

15. *See generally* Nancy Devlin & David Parkin, *Does NICE Have a Cost-Effectiveness Threshold and What other Factors Influence its Decisions? A Binary Choice Analysis*, 13 HEALTH ECON. 437 (2004) (discussing the 20,000–30,000 pound per quality-adjusted life year (QALY) range of acceptable cost-effectiveness that the UK's National Institute for Clinical Excellence uses for its "binary" coverage decisions).

oncology sector, the tradeoff will be even more uncertain, since the benefits are unknown.

Sophisticated economic methods continue to be refined in hopes of somehow bridging the normative-empirical divide, at least when the tangible health benefits can be identified.[16] However, it remains to be seen whether such methods could ever have the reliability, validity, and perceived legitimacy to justify such binary forms of rationing regulation.

GENERALIZED REGULATORY PRESUMPTIONS AS A WAY FORWARD

Instead of trying to ration particular uses of particular technologies for particular patients, regulators and insurers may turn to other mechanisms that work more generally to reduce consumption. For example, policymakers may seek to regulate the ways that the pharmaceutical and medical device industries interact with physicians, proscribing certain sales pitches or economic relationships that drive consumption of medical technologies.[17] Or, policymakers may modify insurance reimbursements in ways that give providers or patients incentives to decline certain medical technologies that they otherwise might consume. In other work, I have suggested novel mechanisms for doing so.[18]

To the extent that these sorts of policy interventions preserve patients' access to and choices about medical technologies, these policy levers may be more feasible in the market and politics than binary rationing policies. The aversion to binary rationing may, moreover, reflect bona fide individual preferences. Individuals may rationally prefer to have access to medical technologies, which they in fact later decline to actually consume. This is the option-value of health insurance.

While these sorts of policies may work to reduce wasteful consumption, they are admittedly crude and are not tailored to individual patients consuming particular treatments for particular purposes (since we lack the normative and epistemic basis for narrower tailoring). The policies will thus reduce consumption of some high-value medical technologies, along with the low-value uses being targeted. Accordingly, there are paternalist concerns that such policies harm the welfare of individual patients.[19] Similarly, there are utilitarian concerns that such reforms may actually setback aggregate social welfare.[20] The problem with these paternalist and welfarist critiques is the same one that

16. *See e.g.*, Emma McIntosh, Cam Donaldson & Mandy Ryan, *Recent Advances in the Methods of Cost-Benefit Analysis in Healthcare*, 15 PHARMACOECONOMICS 357 (1999) (reviewing new methods of assessing benefits, including willingness to pay (WTP) and conjoint analysis, and arguing that the "balance sheet" approach to cost-benefit analysis allows alternative ways of assessing benefit, rather than simple monetization).

17. *See e.g.*, Christopher T. Robertson, *When Truth Cannot Be Presumed: The Regulation of Drug Promotion under an Expanding First Amendment*, 94 B.U. L. REV. 551(2014).

18. *See generally* Robertson *supra* note 5; Christopher T. Robertson, *Scaling Cost-Sharing to Wages*, 14 YALE J. HEALTH POL'Y L. & ETHICS (forthcoming 2014) [hereinafter "Robertson, *Scaling Cost-Sharing to Wages*"].

19. *See* Elhauge, *supra* note 1, at 1480 ("paternalism, perhaps most accurately explains what actually motivates policymakers to refuse to simply redistribute cash: the poor, many believe, would irrationally spend money on the wrong things."); *see also* Robertson, *supra* note 17 (discussing judicial concerns that restrictions on off-label marketing will be deleterious to patients).

20. *See e.g.*, Joseph P. Newhouse, *Medical Care Costs: How Much Welfare Loss?*, 6 J. ECON. PERSPECTIVES 3, 3 (1992) ("some proposed 'cost containment' policies may result in welfare losses for the insured.").

motivates these more generalized policy efforts: we just do not know whether a particular use of a medical technology will promote or set back individual and aggregate social welfare. Such objections can be paralyzing.

There is a way forward. Sometimes belief in general propositions is more warranted than belief in particular propositions, and this may be true even when a particular proposition is a member of the set of propositions that would be covered by the general proposition. Such generalized knowledge may then justify probabilistic knowledge about the specific propositions nonetheless. The courts use such a form of reasoning in *res ipsa loquitor* cases, where specific proof of breach is unavailable, but the plaintiff appeals to a generalized belief that this sort of accident is unlikely to occur in the absence of negligence.[21] When that predicate has been laid, courts then shift the burden to the defendant to give more particularized evidence showing non-negligence.[22]

For a more generalizable example of such reasoning, suppose that I need to decide whether a person named Costa is taller than one named Beneficia. I could invest in a tape measure and hire someone to find Costa and Beneficia and then measure each of them. If successful, such a procedure would give me a particularized knowledge with confidence nearing 100% (barring only fraud and incompetent measurements). But such a measurement procedure may or may not be economically worthwhile, normatively acceptable, or practically feasible. Perhaps Beneficia is hiding in distant country and refuses to be measured, even if found. How else might I proceed? Suppose I also know that Costa is a man and Beneficia is a woman, and have general knowledge that men tend to be taller than women. Even without any further information, I could then infer something about the pairwise relation, a particular proposition that Costa is taller than Beneficia.[23] Of course, the amount of variance, the closeness of the medians, and the direction and size of any skew, would impact the confidence in the proposition that Costa is taller than Beneficia in the distributions.

Now, suppose that I actually measure Costa and learn that he is in the top 5% of all men in terms of height. While that fact tells me nothing about women or about Beneficia in particular, it does tell me something about the particular pairwise comparison: the likelihood that Costa is taller than Beneficia. That fact increases my confidence that Costa is taller than Beneficia, since I have reduced (or perhaps even eliminated) the overlap between the two remaining distributions (the left-censored distribution of the tallest men versus the distribution of all women). With only this knowledge about the underlying distributions of men and women, and Costa's position within the male distribution, I can

21. *See e.g.*, the classic case of *Byrne v. Boadle*, 159 Eng. Rep. 299 (Exch. 1863) (the barrel-flying-out-a-window case).

22. Judge Posner has criticized reliance on a similar form of statistical reasoning, when the proponent "does not show that it was infeasible for him to obtain any additional evidence" that could provide particularized proof. *See* Howard v. Wal–Mart Stores, Inc., 160 F.3d 358 (7th Cir. 1998) (drawing on the famous example of an injury caused by a bus in a town where the defendant owns fifty-one percent of the buses).

23. Not knowing anything about Costa and Beneficia in particular, then I am left to assume that their height is a random variable, with a probability at each possible height given by the known distribution of heights. This inference would be inductive, not deductive. This form of argument is known as the "statistical syllogism," though the use of the term "syllogism" may be controversial, since it may suggest deductive certainty. *See* Carl G. Hempel, *Inductive inconsistencies*, 12 SYNTHESE 439 (1960) (discussing the logical properties of this form of argument).

make a warranted presumption that Costa is in fact taller than Beneficia, with nearly 100% confidence.

One could call this warranted belief "knowledge", though the term may fail to remind users that knowledge is defeasible. Instead, I call this a "presumption" to emphasize that I have still lack particularized evidence about Beneficia, and thus to hold open the possibility that the belief may be falsified if such evidence were produced. Still, given that I have near 100% confidence, I might find it unreasonable to invest in finding and persuading Beneficia to actually measure her. Instead I might invite anyone who seeks to dissuade me to make that investment themselves. In this way, I might allocate a burden of persuasion.

The net value of medical technologies can be analyzed similarly. Rather than comparing male heights to female heights, we instead attempt to determine the costs of medical technologies versus their benefits. For each of these variables, we imagine a distribution including every potential use of every technology. There are highly beneficial uses of technologies and completely useless ones. There are very expensive technologies and extremely cheap ones. Of course each particular cost is linked to a particular benefit, tied by a particular technology used for a particular patient for a particular purpose (similarly you could imagine a height dataset limited to pairs of husbands and wives). The general correlation between these points is unknown, but we are aware of very cheap technologies (e.g., aspirin) that have been shown to have large benefits (e.g., cardioprotection), and very expensive technologies (e.g., Avastin) that have been shown to have no benefit (e.g., for breast cancer survival).[24]

So, to evaluate the cost-effectiveness of any particular technology used for any particular patient with a particular disease at a particular point in time, we could invest in trying to find its place on each distribution, and then compare them. As it happens, it is relatively easy for policymakers to assess the costs of medical technologies, since the producers and providers of these medical technologies put prices upon them and the policymakers pay those bills.[25] Of course, beyond the price, there are other important and sizeable costs that are more difficult to identify and quantify, such as the side effects of taking a drug, and the opportunity cost of taking a new drug versus a standard-of-care drug that may have greater efficacy. But actual prices paid provide a readily accessible lower bound on the cost of a treatment.

Like the supposed difficulty of measuring Beneficia's height, the quantification task is profoundly difficult for the benefits of medical technologies, for the reasons Elhauge identifies.[26] Nonetheless, we may be able to make generalized claims about how the two distributions are shaped and related, which will allow less confident but still warranted inferences about likely pairwise comparisons, between known costs and likely benefits. As such, even while lacking particularized information about benefits of particular medical technologies used for individual patients for specific treatments, this form of inference allows us to assess the likelihood that a policy intervention that attempts to

24. *See* discussion *infra* notes 53-58 and accompanying text.

25. Frustratingly, these prices are not often accessible to patients. *See* Uwe E. Reinhardt, *The Pricing of U.S. Hospital Services: Chaos Behind a Veil of Secrecy*, 25 HEALTH AFF. 57, 58-62 (2006).

26. *See supra* Part I.

reduce consumption of these most expensive treatments will overall harm individual patients or setback aggregate social welfare. We can make such inferences even if we know absolutely nothing about the benefits of a particular use of a particular medical technology.

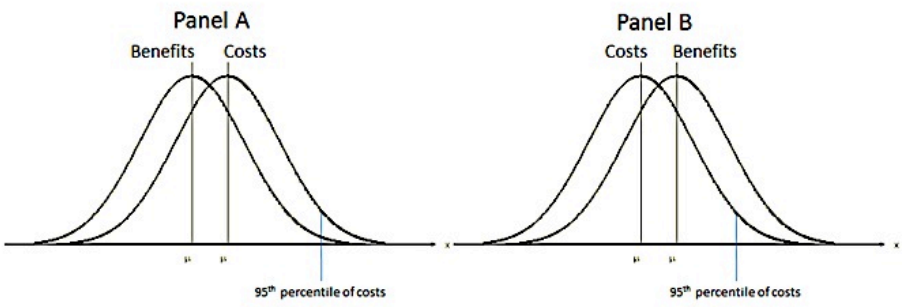
Our inferences will be stronger if we limit our comparisons to the tail of one of the distributions, as we did by measuring Costa and finding that he was one of the tallest men. In this domain it is quite effective for policymakers to target the top 5% of medical technologies by price (i.e., the 95th percentile), since healthcare spending is very highly concentrated. In a given year, the top 5% of individuals (who spend about $50,000 on average on healthcare) account for half of all healthcare spending.[27] So, if policymakers somehow discouraged consumption of the most expensive medical technologies, they could make a sizeable impact on aggregate healthcare spending.

Using this method, as shown in Figure 1 Panel A, we can make confident inferences that uses of medical technologies in that sector of the top 5% of costliest technologies are unlikely to have exceeding benefits. Like height for any pair of men and women, this Figure supposes a common scale (e.g., dollars) for assessing costs and benefits, but there are two distributions because the two amounts may not be correlated for any particular drug. Still, it must be conceded that we do not know the shape or position of the distribution of benefits in the way that we know the distribution of heights for women. Thus our conclusions in the domain of costs and benefits will be more tentative, than our conclusions about height.

It is possible that the costs distribution is shifted to the right of the benefits distribution, as shown in Panel A. Alternatively, Panel B concedes, *arguendo*, that the benefits may generally outweigh the costs (depicted by its shift to the right). Even then, Panel B shows how the top 5% of most expensive drugs are likely to have costs that are greater than the likely benefits. This would be analogous to the claim that a woman who is in the top 5% of all women's heights is taller than a man with unknown height.

27. *See* Steven B. Cohen & William Yu, *The Concentration and Persistence in the Level of Health Expenditures over Time: Estimates for the U.S. Population, 2008–2009*, MED. EXPENDITURE PANEL SURV. (Jan. 2012), http://meps.ahrq.gov/mepsweb/data_files/publications/st354/stat354.pdf.

Figure 1—Hypothesized distributions of costs and benefits of medical technologies.



Note: Any particular use of a product will have a point on each distribution, and cost-benefit analysis seeks to compare them. To a policymaker, the cost is often known, but particular benefit often unknown, and thus the propensity for being at any point on the benefit distribution is averaged. The plot illustrates the unlikelihood that the known costs of any one of the top 5% of most expensive technologies will exceed its unknown benefits. This presumption holds even if one assumes that benefits generally exceed costs (i.e., the right-shift of the benefits distribution shown in Panel B). A greater shift or a strong correlation between costs and benefits across products would undermine this presumption.

WHY A PRESUMPTION AGAINST CONSUMPTION IS WARRANTED

Now that the strategy for defending a presumption against consumption has been laid bare, it remains incumbent to provide the predicate facts about the cost and benefit distributions, which would support the inference that the costs of the most expensive medical technologies are likely to exceed their benefits. Let me specify some falsification criteria for this generalized probability thesis, which will then become the targets of the remainder of this essay. Belief would be unwarranted if, as a matter of fact, the aggregate distribution of the benefits of particular uses of medical technologies were on the whole much greater than the aggregate distribution of costs. In this world, the right-shift hypothesized in Figure 2, Panel B would be even greater than that shown, thereby reducing overlap. Similarly, if the benefits were right-skewed or the costs left-skewed (rather than the normal distributions shown in both Panels of Figure 2), the presumption against consumption would be unwarranted. Finally, if there were a strong correlation between costs and benefits for individual uses of medical technologies (so that the most expensive ones also tended to have the greatest benefits, so that referring to the top 5% of costs would correlate with the top 5% of benefits), then this hypothesis would be falsified.[28]

28. There is some evidence produced by the pharmaceutical industry that prices do track benefits. *See* Z. John Lu & William S. Comanor, *Strategic Pricing of New Pharmaceuticals*, 80 REV. OF ECON. & STAT.

Although I have tried to be quite precise about how such generalized arguments can generate probabilistic knowledge, the remainder of the essay will draw on various sorts of indirect, imprecise, and informal evidence to characterize in a more synthetic way the cost and benefit distributions and the relationship between them. Ultimately, I need only show that the relationship shown in either of the Panels in Figure 2 is about right. There are a dozen reasons to suspect that the falsification criteria are unmet, including: the disconnect between research investments and health outcomes, the FDA's lack of authority to consider cost-effectiveness, and the prevalence of off-label uses and the lack of supporting evidence therefore. There are also problems in scientific research including publication biases, lack of effective randomization and blinding, commercial biases, the use of proxies for efficacy, small demonstrated benefits not enjoyed by most consumers, and the lack of power to detect adverse outcomes. There are also market failures because consumers are unable to perceive value and have little or no exposure to cost, while their advisors, physicians, have misaligned incentives. Ultimately, aggregate data across time, geography, and experimental conditions shows that much medical spending is on the "flat of the curve," not delivering commensurate healthcare value.

First, let us begin with the most common refrain: "[e]very time there is a public debate about drug prices, the pharmaceutical industry replies . . . [that] [t]he cost of bringing a new drug to market is enormous—$1.3 billion per drug, according to one often-cited (but often-contested) academic study."[29] A similar story could be told for other expensive medical technologies, such as devices. Prior research has shown that very little of industry spending is actually directed towards the discovery of new products, and much more is directed to marketing.[30] Regardless, one can concede the high costs and grant that industry may rationally seek to defray such costs by charging high prices, but those concessions say nothing about whether it is rational for anyone to consume the product at the demanded price. That question, instead, turns on the likely benefits compared to available alternatives.[31] Likewise, an explorer may invest a fortune to find the world's largest emerald, but once the emerald is found, its value will be determined by comparing its size and quality to the next best emerald. Its price cannot be justified by the sunk costs to find it. The United States government already subsidizes the development of medical technologies in many ways; purchasing high-cost low-value products is a poor way to provide a subsidy, since it incentivizes the wrong products. Thus, the consumption question must remain focused on the relationship of costs to payers and benefits to patients, at the point of consumption.

Second, it would be attractive for policymakers to defer to decisions of an expert agency, such as the FDA, as a proxy for determining whether consumption of medical technologies is worthwhile. Accordingly, one might suppose that the mere fact that an

108 (1998).

29. Hall, *supra* note 5, at 3..

30. Donald W. Light & Joel R. Lexchin, *Pharmaceutical Research and Development—What Do We Get for All That Money?* BMJ, at 2 (2012), www.pharmamyths.net/files/BMJ-Innova_ARTICLE_8-11-12.pdf (finding that the pharmaceutical industry spends only 1.3 percent of revenues, net of taxpayer subsidies, to discovering new molecules, and that they spend nineteen times more than that on promotion).

31. *See* F.M. Scherer, *The Pharmaceutical Industry—Prices and Progress*, 351 NEW ENG. J. MED. 927, 928-29 (2004) (debunking this "fallacy" of "sunk costs").

expensive medical technology has entered the market, garnering the approval of the FDA, suggests that its benefits exceed its costs. That supposition would be unfounded. Initially, it is worth noting the real possibility that agency capture has corrupted the agency's decisions on the margins, causing it to err on the side of approving drugs that may actually have more dangers than benefits.[32] Even if the FDA were perfectly functioning, the more fundamental problem is that the FDA statute does not authorize it to consider costs at all.[33] The FDA requires only that manufacturers prove minimal effectiveness compared to a placebo, which is to say that the product is better than nothing.[34] Even when proven efficacious compared to a placebo, there are often no studies comparing the new medical technologies to cheaper, standard treatments, and thus we cannot say whether they have marginal benefits to justify their higher costs.[35] And for medical devices, the FDA's scrutiny is even lighter. For example, for seventy-eight "high-risk" cardiovascular devices that the FDA approved between 2000 and 2007, less than one third had been subjected to a randomized trial, and only 5% had undergone two or more blinded randomized studies.[36] Once a device has been approved, manufacturers can then get FDA approval for newer versions of the device, without data to prove that the new design is safe or effective.[37]

Even more, the FDA statute also allows physicians to prescribe drugs and devices "off-label" for other, unapproved diseases and conditions without any proof or FDA review of efficacy.[38] Thus, for many uses of medical technology, there is no FDA review

32. *See* Donald W. Light, Joel Lexchin & Jonathan J. Darrow, *Institutional Corruption of Pharmaceuticals and the Myth of Safe and Effective Drugs*, 14 J. L. MED. & ETHICS 590, 597 (2013) (discussing "a 15-month investigation by the Committee on Government Reform of the U.S. House of Representatives found '[a] growing laxity in FDA's surveillance and enforcement procedures, a dangerous decline in regulatory vigilance, and an obvious unwillingness to move forward even on claims from its own field offices.'" *See id.* (quoting Henry A. Waxman, *Prescription for Harm: The Decline in FDA Enforcement Activity*, COMM. ON OVERSIGHT AND GOV'T REFORM (June 26, 2006), http://oversight-archive.waxman.house.gov/story.asp?ID=1074)).

33. *See* Peter J. Neumann et al., *The FDA and Regulation of Cost-Effectiveness Claims*, 15 HEALTH AFFS 54, 55-56, 68 (1996). Still, the insurers' actuaries may have developed such data internally, and recent public policy initiatives should expand our knowledge in this area. *See Recovery Act Allocates $1.1 Billion for Comparative Effectiveness Research*, U.S. DEP'T HEALTH & HUM. SERVS., http://wayback.archive-it.org/3909/20130927160907/http://www.hhs.gov/recovery/programs/os/cerbios.html (last visited Feb. 28, 2014).

34. *See* Jerry Avorn, *FDA Standards—Good Enough for Government Work?*, 353 NEW ENG. J. MED. 969, 969 (2005). *See generally* Riegel v. Medtronic, Inc., 552 U.S. 312, 316 (2008) (analyzing the preemptive effect of the U.S. Food and Drug Administration ("FDA") statute for medical devices); United States v. Rutherford, 442 U.S. 544, 555 (1979) (discussing the purpose and function of the FDA approval process and upholding a ban on the sale of unapproved drugs).

35. *See generally* Agency Information Collection Activities; Proposed Collection; Comment Request; Experimental Study of Comparative Direct-to-Consumer Advertising, 76 Fed. Reg. 38663-01, 38664 (July 1, 2011) (noting that "few head-to-head clinical trials have been conducted"). *See* Nicholas S. Downing, et al., Clinical Trial Evidence Supporting FDA Approval of Novel Therapeutic Agents, 2005-2012, 311 J. AM. MED. ASS'N 368, 374 (2014) ("comparison of an intervention with an active control was available for less than half of indications.").

36. *See generally* Sanket S. Dhruva, Lisa A. Bero & Rita F. Redberg, *Strength of Study Evidence Examined by the FDA in Premarket Approval of Cardiovascular Devices*, 302 J. AM. MED. ASS'N. 2679 (2009).

37. *See* Benjamin N. Rome, Daniel B. Kramer & Aaron S. Kesselheim, *FDA Approval of Cardiac Implantable Electronic Devices via Original and Supplement Premarket Approval Pathways, 1979-2012*, 311 J. AM. MED. ASS'N 385 (2014).

38. *See* 21 U.S.C. § 396 (2009) ("Nothing in this [Act] shall be construed to limit or interfere with the authority of a health care practitioner to prescribe or administer any legally marketed device to a patient for any condition or disease within a legitimate health care practitioner-patient relationship."); Buckman

of the efficacy and risks of physician's intended use. The FDA thus cannot reassure that the benefits distribution is far to the right of the costs distribution.

Third, one might suppose that scientific testing has shown important benefits for expensive medical technologies. Yet, many uses of medical technologies are altogether lacking in any evidentiary basis. A 2006 study found that twenty-one percent of all prescriptions written in the United States are for off-label uses, and that most of these had "little or no scientific support."[39] Off-label, unsupported use is particularly common in high-cost domains such as oncology and cardiology. For example, surgeons implant more than one million heart stents each year, at a cost of $30,000 to $100,000 each.[40] But, more often than not, these stents are used off-label, in settings where they have not been proven at all effective.[41] Indeed, a large randomized, controlled trial demonstrated that patients who received stents off-label would have fared just as well on a much cheaper (and safer) regimen of drugs.[42] More broadly, the Congressional Budget Office ("CBO") has concluded that, "[a]lthough estimates vary, some experts believe that less than half of all medical care is based on or supported by adequate evidence about its effectiveness."[43] A recent review of every article published in the *New England Journal of Medicine* for a decade found that when randomized trials were utilized to test the efficacy of commonly used medical procedures, they failed more often than passed.[44]

Fourth, concededly, there is a domain of healthcare consumption that *is* supported by the peer reviewed scientific literature, and in this domain one might have greater hope that the benefits of a medical technology will exceed the costs, since benefits are at least proven to exist. Yet, these estimates of efficacy have significant problems.

Publication bias is rampant. Due to random variation (including sampling error), some studies will overestimate the true effects and other studies will underestimate the true effects. Confidence intervals and p-values are supposed to help readers differentiate between spurious and real effects. But there is a larger selection bias in the known studies. When a scientific study shows that a medical technology is not effective, the funders and authors may decline to submit it for publication. Even if submitted, for studies of

---

Co. v. Plaintiffs' Legal Comm., 531 U.S. 341, 350–51 (2001) (discussing off-label use in the context of medical devices).

39. David C. Radley et al., *Off-Label Prescribing Among Office-Based Physicians*, 166 ARCHIVES INTERNAL MED. 1021, 1021 (2006).

40. While the cost of a stent can vary between $1,000 and $4,000, the average cost of a stent procedure can vary from $30,000 to $100,000. *See* David Rosenfeld, *Is American Medicine Too Stent Happy?*, PAC. STANDARD (Apr. 17, 2010), http://psmag.com/health/is-american-medicine-too-stent-happy-12861.

41. Htut K. Win et al., *Clinical Outcomes and Stent Thrombosis Following Off-Label Use of Drug-Eluting Stents*, 297 J. AM. MED. ASS'N. 2001, 2008 (2007) (finding that of the 3,323 patients enrolled in the study who had received stents, 54.7 percent had at least one off-label characteristic).

42. *See* William E. Boden et al., *Optimal Medical Therapy with or without PCI for Stable Coronary Disease*, 356 NEW ENG. J. MED. 1503, 1503 (2007).

43. Cong. Budget Office, *Research on the Comparative Effectiveness of Medical Treatments: Issues and Options for an Expanded Federal Role* 11 (2007) [hereinafter Cong. Budget Office, *Research on the Comparative Effectiveness of Medical Treatments*], *available at* http://www.cbo.gov/ftpdocs/88xx/doc8891/MainText.3.1.shtml (citing Leigh Anne Olsen et al., *Learning What Works Best: The Nation's Need for Evidence on Comparative Effectiveness in Health Care*, INST. OF MED. OF THE NAT'L ACADS. 341 (2011), *available at* http://www.ncbi.nlm.nih.gov/books/NBK64787/pdf/TOC.pdf).

44. Vinay Prasad et al., *A Decade of Reversal: An Analysis of 146 Contradicted Medical Practices*, 88 MAYO CLINIC PROCEEDINGS 790, 792 (2013).

new drugs and devices that are not yet widely used, journal editors tend to prefer studies that show significant results. Thus, for a physician or policymaker to evaluate the efficacy of a medical technology one must imagine a bell-shaped distribution of scientific results, from which we only get to see a hand-picked sample of those that purport to show the greatest efficacy. This means that published research systematically overestimates effect size.[45] When publication bias is taken into account, the proven efficacy of medical technologies sometimes disappears altogether.[46] Unfortunately, these sorts of adjustments are not routinely made.

Fifth, medical efficacy tends to be overestimated for another reason: lack of random assignment and either lack of blinding or ineffective blinding, in the underlying studies. The gold standard for scientific research is the double-blinded randomized placebo-controlled study, since it prevents the participants, the treaters, and the raters from self-selecting into the treatment arm of the study and from then overestimating the benefits of the treatment. Yet, a huge portion of the scientific literature is based on un-blinded non-randomized studies, rather than in randomized placebo-controlled trials. This dynamic is especially common for exploring the claimed efficacy of medical devices, where it seems infeasible to undertake sham surgeries to implant such devices.[47] In a 2012 study of trials in a variety of disease areas that had both blinded and non-blinded outcome assessors, Hróbjartsson and colleagues estimated the size of the bias due to lack of blinding, and found a median shift in the odds ratio towards efficacy by about 36 %.[48] Disconcertingly, this shift was larger than the proven effect size for most of the treatments tested. Here again efficacy disappears when bias is properly accounted for.

Sixth, the biomedical scientific literature also overestimates efficacy because the drug and device industry is itself the funder and designers of the vast majority of the scientific trials of its own products.[49] This funding mechanism means that the drug and device industry sets the scientific agenda, de-prioritizing the study of unpatented medical technologies and other non-technology interventions, which may deliver greater value. Even for the studies the industry does run, Judge Jack Weinstein writes that, "The pervasive commercial bias found in today's research laboratories means studies are often lacking in essential objectivity, with the potential for misinformation, skewed results, or cover-ups."[50] Similarly, an Institute of Medicine report concluded that "[s]everal systematic reviews and other studies provide substantial evidence that clinical trials with industry ties are more likely to have results that favor industry."[51] It is possible, of course, that

45. David M. Lane & William P. Dunlap, *Estimating effect size: Bias resulting from the significance criterion in editorial decisions*, 31 BRIT. J. MATHEMATICAL STATISTICAL PSYCHOL. 107 (1978).

46. *See* Sue Duvaland & Richard Tweedie, *Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis*, 56 BIOMETRICS 455, 461 (2000) (showing such an example for anti-malaria drugs).

47. *See* Dhruva et al., *supra* note 36.

48. Asbjørn Hróbjartsson et al., *Observer bias in randomized clinical trials with binary outcomes: Systematic review of trials with both blinded and non-blinded assessors*, 333 BRIT. MED. J. 231 (2012).

49. *See generally* Christopher T. Robertson, *The Money Blind: How to Stop Industry Bias in Biomedical Science, Without Violating the First Amendment*, 37 AM. J.L. & MED. 358, 362 (2011) (reviewing this literature and listing eighteen ways in which industry can bias the design, conduct, and reporting of studies).

50. *In re* Zyprexa Prods. Liab. Litig., 253 F.R.D. 69, 106 (E.D.N.Y. 2008), *rev'd sub nom.* UFCW Local 1776 v. Eli Lilly & Co., 620 F.3d 121, 133 (2d Cir. 2010) (reversing on the question of causation).

51. *Conflict of Interest in Medical Research, Education, and Practice*, INST. OF MED., 104 (Bernard Lo &

independently-funded trials actually underestimate the benefits, but it seems more plausible to suppose that industry trials over-estimate the benefits, due to self-interested decision-making about how to design, conduct, analyze, discuss, and publish the research study.

Seventh, some drugs are "proven effective" only because they promote certain outcomes that are thought to be surrogates for mortality and morbidity, even if those health outcomes cannot be proven directly.[52] Thus proof of efficacy may not actually mean proof of any real benefit to patients. Avastin, for example, costs $88,000 per treatment,[53] and received accelerated approval for breast cancer in 2008.[54] The approval was based on two unblinded studies that showed reduced tumor growth but "no evidence of an effect on overall survival or improved symptoms."[55] Subsequent double-blinded studies failed to replicate the early findings.[56] Notwithstanding continued widespread use of the drug by oncologists, in 2011, the FDA commissioner revoked Avastin's approved labeling as a treatment for breast cancer, noting that its side effects presented very real dangers to patients (including a risk of death) not balanced by a proven health benefit.[57] The drug continues to be prescribed and reimbursed by public and private insurers.

Eighth, even after one adjusts for publication bias, unblinding, and industry influence, and finds proof of efficacy for a real health outcome, it must be understood in relative terms. Efficacy is only proven on the margin, and those marginal benefits may be quite small in the real world.[58] Even with an important outcome like survival, the marginal benefits may be small—measured in terms of weeks or months.[59] Indeed, many, or even most of those using the medical technology will get no benefit whatsoever.[60] Imagine a drug that shifts two-month survival from 10% of patients on placebo to 20% of patients taking the drug. That doubling of the survival rate would be a sizeable effect, and could be expressed as a large odds ratio with statistical significance. However, that outcome also means that for every ten patients that consume the drug, eight would have died regardless, and one would have survived regardless. Only one out of ten received a benefit on the margin; the remaining 90% paid the costs and suffered the side effects, without

Marilyn J. Field eds., 2009).

52. *See* Downing et al., *supra* note 35, at 374 ("Pivotal trials using surrogate end points as their primary outcome formed the exclusive basis of approval for 91 indications (45.3%).").

53. Merrill Goozner, *An Extra Month of Life: Is It Really Worth the Cost?*, FISCAL TIMES, June 30, 2011, http://www.thefiscaltimes.com/Articles/2011/06/30/An-Extra-Month-of-Life-Is-it-Really-Worth-the-Cost.aspx#page1.

54. *See* Proposal to Withdraw Approval for the Breast Cancer Indication for AVASTIN (Bevacizumab), Docket No. FDA-2010-N-0621, at 23 (Dep't of Health & Human Servs., Food & Drug Admin. Nov. 18, 2011) [hereinafter Decision of the Commissioner].

55. *Id.* at 13. The studies submitted to obtain accelerated approval (the E2100 study and the AVF2119g study) considered tumor growth to constitute disease progression (PFS). *Id.* at 18.

56. *See id.* at 23–25.

57. *Id.* at 12, 40.

58. It is important to distinguish between statistical significance and effect size. *See* Jonathan J. Darrow, *Crowdsourcing Clinical Trials,* 98 MINN. LAW REV (forthcoming 2014) ("The absence of substantial efficacy for newly FDA-approved drugs is entirely legal under United States law, which requires only that the *evidence* of efficacy be substantial.").

59. *See* Abboud, *supra* note 14 (discussing the "minor" benefits of cancer drugs).

60. *See* Amitabh Chandra & Jonathan Skinner, *Technology Growth and Expenditure Growth in Health Care*, 50 J. ECON. LIT. 645-80 (2012) (describing a broad swath of medical technologies that have heterogeneous effects across patients).

any benefit. This concept can be expressed by calculating the "number needed to treat" ("NNT") one patient.[61] It is quite common for medical technologies to have NNT scores above five, which means that less than one out of every five patients will benefit, and quite rare to have an NNT less than two, which would mean that most patients receive a benefit from consuming the medical technology.[62]

Ninth, the most expensive medical technologies are often also the newest ones on the market, since they purport to be an advance on older treatments.[63] And the newest ones are the ones that the FDA, scientists, and physicians know the least about. Clinical trials are typically designed with enough statistical power and a long enough period of follow-up to detect potential clinical benefits, but are not able to detect rare but severe adverse events, like death (which must be considered in any cost-benefit analysis). The median "pivotal" clinical trial that the FDA uses to approve a drug has 446 patients taking the drug for about 14 weeks.[64] Jonathan Darrow has thus argued that prescribing of newly released drugs is tantamount to human subjects research, without the informed consent of the patients.[65] It is hard to forget Vioxx (rofecoxib), the pain medicine launched that is estimated to have caused the deaths of nearly 28,000 people in the five years before it was removed from the market.[66] About one in every twenty-five new drugs is removed from the market as being simply too dangerous.[67] Most of these recalls happen within the first few years on the market.[68] It is ironic that older, cheaper drugs that have survived this gauntlet thus tend to be safer than the newer, expensive drugs, whose risk profile is poorly understood.

Tenth, one must consider whether "[t]he market, in its own way, provides information about individual preferences. When people decide to buy or not to buy, or to offer for sale or not offer for sale, they indicate what things are worth to them."[69] On this perspective—something like the efficient market hypothesis proposed to understand the stock market—the consumer is presumed to be making value-enhancing trades between money and healthcare. In reality, we know that the market for medical technologies is far

61. Emmanuel Lesaffre, *Number Needed to Treat*, 3 ENCYCLOPEDIA OF STATISTICS IN BEHAVIORAL SCI. 1448 (2005) (discussing advantages and disadvantages of using number needed to treat).

62. *See Therapy NNT Reviews*, THE NNT, *available at* http://www.thennt.com/home-nnt/.

63. Abboud, *supra* note 14 (answering the question of how cancer drugs are priced by stating "[o]f the many complex factors involved, price often seems to follow a simple formula: start with the price for the most recent similar drug on the market and price the new one within 10-20% of that price (usually higher)."). *See also* Peter B. Bach, *Limits on Medicare's Ability to Control Rising Spending on Cancer Drugs*, 360 NEW ENG. J. MED. 626 (2009) (plotting the dramatic increase in new drug prices over time). Furthermore, within the market-life of a drug, prices will tend to go down as substitution drugs come onto the market and the patent expires.

64. Downing, et al., *supra* note 35, at 371.

65. *See generally* Darrow, *supra* note 58.

66. *FDA, Merck, and Vioxx: Putting Patient Safety First?: Hearing Before the S. Comm. on Fin.*, 108th Cong. 14, at 13 (2004) (statement of David J. Graham, M.D., M.P.H., Associate Director for Science, Office of Drug Safety, Food and Drug Administration) (describing the 28,000 figure as "an extremely conservative estimate").

67. *See* Zaina P. Qureshi et al., *Market Withdrawal of New Molecular Entities Approved in the United States from 1980 to 2009*, at 20 PHARMACOEPIDEMIOLOGY & DRUG SAFETY 772, 775 (2011) (showing thirty out of 740 new drugs withdrawn); *see also* Amalia M. Issa et al., *Drug Withdrawals in the United States: A Systematic Review of the Evidence and Analysis of Trends*, 2 CURRENT DRUG SAFETY 177, 179 (2007) (similar for twenty drugs withdrawn between 1993 and 2006).

68. Darrow, *supra* note 58.

69. Mark V. Pauly, *Indemnity Insurance for Health Care Efficiency*, 24 ECON. & BUS. BULL. 53, 53 (1971).

from efficient. Public or private insurance covers most of the consumption choices, which creates "moral hazard," a problem where the decision maker internalizes the benefits but externalizes the costs.[70] The largest such insurer, Medicare, is prohibited by law from negotiating with the providers of medical technologies, and the "absolutist imperative" to cover all "medically necessary care" prevents insurers from negotiating very much.[71] Copays do cause patients to consider a small fraction of the costs, which they must pay out of pocket. However, copays are largely irrelevant in the domain of the highest cost medical technologies, since:

> Roughly 77% of full-time employees of medium and large establishments enrolled in non-HMO plans have maximum out-of-pocket limits less than US$2000 per individual and the most common coinsurance rate is 20% . . . . Thus, individuals with more than US$10,000 in total costs will face no cost sharing at the margin.[72]

As Timothy Jost explains, "[o]nce consumers reach the limits of the deductible, they have little reason to limit their consumption of health care or to pay attention to its price."[73]

While patients are largely insulated from the costs of medical technologies, they are also poor judges of the benefits. In one study of actual patients receiving chemotherapy for their terminal lung and colorectal cancers, the vast majority of them were under the misconception that the chemotherapy could be curative.[74] Other studies have shown that most consumers overestimate drug efficacy by a factor of ten or more.[75] Such rampant inaccuracies may be due to optimism bias, combined with the difficulty of understanding highly technical information.[76]

Eleventh, one might suppose that the benefits of medical technologies will exceed their costs because physicians stand as expert gatekeepers, in a distributed system of regulation. While one might have reason to hope that physicians do not make such severe errors in estimating benefits as patients, it bears emphasis that the scientific basis for

70. *See National Health Expenditures By Type Of Service And Source Of Funds,* Ctrs. For Medicare & Medicaid Servs., CY1960–2011, *available at* http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html (showing that public and private insurers pay over eighty-eight percent of health care costs).

71. *See generally* Bach, *supra* note 63; *see also supra* notes 8-16 and accompanying text (discussing the absolutist imperative).

72. Michael E. Chernew et al., *Optimal Health Insurance: The Case of Observable, Severe Illness*, 19 J. HEALTH ECON. 585, 588 (2000).

73. Timothy Stoltzfus Jost, *Our Broken Health Care System and How to Fix It: An Essay on Health Law and Policy*, 41 WAKE FOREST L. REV. 537, 587 (2006).

74. Jane C. Weeks, *Patients' Expectations about Effects of Chemotherapy for Advanced Cancer*, 367 NEW ENG. J. MED. 1616, 1619 (2012).

75. Lisa M. Schwartz et al., *Using a Drug Fact Box to Communicate Drug Benefits and Harms*, 150 ANNALS OF INTERNAL MED. 516, 524 (2009).

76. *See* Ken Murray, *Why Doctors Die Differently*, WALL STREET J., Feb 25, 2012, *available at* http://online.wsj.com/news/articles/SB10001424052970203918304577243321242833962 (discussing several studies showing that physicians have different preferences for end of life care than the general public, perhaps because the efficacy of care is misportrayed in mass media).

their decisions is severely constrained, as discussed in points #2-#9 above. Further, the physicians are largely insulated from costs, and cannot be taken to be making cost-benefit tradeoffs. If anything, physician decisions are biased towards costly procedures, given our fee-for-service healthcare finance system, physician self-referrals, and financial relationships with industry, which have all been shown to bias professional judgment.[77] In high-costs fields like oncology and surgery, the incentives to provide high-cost care are particularly stark, since the physician who recommends the care will often be paid to provide it.[78] In fields like cardiology, hematology, oncology, and urology, about half of the physicians also take money directly from the drug and device industries.[79]

Twelfth and finally, one might look to the aggregate data about health spending, to see whether more of it drives beneficial health outcomes. For this purpose, one can examine geographic variation across countries and across regions within the United States, variations across time, and variations between conditions in experimental research.

Geographically, the United States spends twenty to thirty percent more per capita than countries with excellent health care systems and similarly aging and equally or more healthy populations, such as France and Germany.[80] As Ezekiel Emanuel writes, "[t]he truth is, the United States is not getting 20 or 30 percent better health care or results than other countries."[81] Within the United States, from one region to another, there are also large disparities in the amount spent on health care, which cannot be explained by demographic factors and which do not seem to result in improved health.[82] Areas with double, or even triple, the amount of spending per patient do not show better outcomes as a result.[83] This phenomenon has been called "flat of the curve medicine," because increases in spending do not lead to improved health outcomes.[84]

Across several decades, Peter Bach has compiled and plotted the prices of new

77. *See* Christopher T. Robertson, Susannah Rose & Aaron S. Kesselheim, *Effect of Financial Relationships on the Behaviors of Health Care Professionals: A Review of the Evidence*, 40 J.L. MED. & ETHICS 452, 458 (2012); *see generally* Sunita Sah & Adriane Fugh-Berman, *Physicians under the Influence: Social Psychology and Industry Marketing Strategies,* 41 J.L. MED. & ETHICS 665 (2013).

78. *See generally* Mireille Jacobson et al., *How Medicare's Payment Cuts for Cancer Chemotherapy Drugs Changed Patterns of Treatment*, 29 HEALTH AFFS. 1391, 1391 (2010) (assessing how "substantially reduced payment rates for chemotherapy drugs . . . affected the likelihood and setting of chemotherapy treatment for Medicare beneficiaries"); Jean M. Mitchell, *Do Financial Incentives Linked to Ownership of Specialty Hospitals Affect Physicians' Practice Patterns?*, 46 MED. CARE 732, 736 (2008).

79. Aaron S. Kesselheim, Christopher T. Robertson, Kent Siri, Puneet Batra & Jessica Franklin, *Distributions of Industry Payments to Massachusetts Physicians*, NEW ENG. J. MED. (2013), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2261291.

80. *See* Ezekiel J. Emanuel, *How Much Does Health Cost?*, N.Y. TIMES, Oct. 30, 2011, http://query.nytimes.com/gst/fullpage.html?res=9E06E7D9143DF933A05753C1A9679D8B63..

81. *Id.*

82. *See id.*

83. *See id.*; *but see* Richard A. Cooper, *States with More Health Care Spending Have Better-Quality Health Care: Lessons about Medicare*, 28 HEALTH AFFS. w103, w112–13 (2009) (noting that quality "depends on total health care spending" and "relates to a broad array of sociodemographic characteristics," and that "Medicare spending is a poor proxy for overall health care spending.").

84. *See e.g.*, Victor R. Fuchs, *More Variation in Use of Care, More Flat-of-the-Curve Medicine*, HEALTH AFFS. (October 2004), http://content.healthaffairs.org/content/early/2004/10/07/hlthaff.var.104.full.pdf+html (discussing Alain Enthoven as having popularized the term).

cancer drugs upon their date of market entry, adjusting for inflation.[85] In real dollars, new drugs are hundreds of times more expensive than new drugs were in the 1980s.[86] More specifically, Deborah Schrag has examined the important changes in the treatments available for colorectal cancer that emerged between the early 1990s and 2002. Although still failing to cure for the disease, the new drug therapies nearly doubled the median duration of survival, from twelve months to twenty-one months.[87] This is a success. However, Schrag explains that, "[t]he near-doubling of the median survival achieved over the past decade has been accompanied by a staggering 340-fold increase in drug costs—just for the initial eight weeks of treatment."[88] Without purporting to put a price on those additional nine months of survival, one can observe that the growth lines of cost and benefits are not parallel, and thus get a sense that the distribution of overall costs may be moving to the right of overall benefits. Likewise, consider aggregate data from Canada in a decade when spending on drugs more than doubled.[89] A regulatory board appraised the value of 1,147 new drugs introduced during that time: "Of these new drugs, 68 (5.9%) met the regulatory criterion of being a breakthrough drug ('the first drug to treat effectively a particular illness or which provides a substantial improvement over existing drug products') . . . . The remaining 1005 new drugs did not provide a 'substantial improvement over existing drug products.'"[90] The authors concluded that "most (80%) of the increase in drug expenditure between 1996 and 2003 was explained by the use of new, patented drug products that did not offer substantial improvements on less expensive alternatives available before 1990."[91]

Field experiments tell a similar story. In the well-known RAND Health Insurance Experiment, individuals were randomized into conditions with larger or smaller cost-sharing obligations. The research found that larger cost-sharing obligations reduced health spending, without negatively impacting the health of the median participant.[92] Another randomized experiment assigned a group to receive Medicaid benefits and "showed that Medicaid coverage generated no significant improvements in measured physical health outcomes in the first 2 years, but it did increase the use of health services[.]"[93] Although these studies are not focused on the most expensive instances of healthcare consumption, they do undermine any conception that more healthcare consumption necessarily yields more health.

Together these macro-level observations across geography, across time, and in

85. *See* Bach, *supra* note 63.

86. *See id.*

87. Deborah Schrag, *The Price Tag on Progress—Chemotherapy for Colorectal Cancer*, 351 NEW ENG. J. MED. 317, 317-19 (2004).

88. *Id.* at 318.

89. Steven G. Morgan et al., *"Breakthrough" Drugs and Growth in Expenditure on Prescription Drugs in Canada*, BMJ (July 22, 2005), www.bmj.com/content/331/7520/815.

90. *Id. See also* JERRY AVORN, POWERFUL MEDICINES: THE BENEFITS, RISKS, AND COSTS OF PRESCRIPTION DRUGS 198-216 (2004) (describing similar results from an American study).

91. Morgan et al., *supra* note 89.

92. *See generally* JOSEPH P. NEWHOUSE, FREE FOR ALL?: LESSONS FROM THE RAND HEALTH INSURANCE EXPERIMENT (1993).

93. Katherine Baicker et al., *The Oregon Experiment—Effects of Medicaid on Clinical Outcomes*, 368 NEW ENG. J. MED. 1713 (2013). Medicaid coverage did "raise rates of diabetes detection and management, lower rates of depression, and reduce financial strain." *Id.*

field experiments suggest that the cost distribution for medical technologies may actually be shifted to the right of the benefits distribution (as in Panel A of Figure 2), since we are seeing marginal increases in costs not leading to improvements in health benefits. Still, it is possible that some costly procedures provide other non-health benefits (e.g., pain relief) that are lost in these large-scale observational studies, which focus on observable mortality and morbidity statistics.[94]

CONCLUSIONS

Since we have not yet resolved the profoundly difficult questions about rationing that Elhauge identified twenty years ago, we remain stuck in the absolutist imperative to provide patients with access to expensive medical technologies that may have little or no real benefits. Instead of using such overt binary rationing mechanisms, however, we can adopt, as a matter of regulatory policy and insurance design, a generalized presumption against healthcare consumption, at least in the domain of the most expensive outliers. The dozen considerations limned herein make it seem unlikely that the benefits of the most expensive medical technologies will usually exceed their costs. Policymakers can use incentives, nudges, and other policies to discourage consumption of, even while preserving choice and access to, these technologies. Such defeasible policies thereby remain open to the inevitable counter-examples of expensive treatments that are highly beneficial, when proponents can meet that burden of particularized persuasion.[95]

The foregoing considerations have mostly focused on the question of whether spending on expensive medical technologies is likely to produce enough health to justify their costs. But of course health is not our only goal, and medical technologies are not the only means to that end. As Professor Elhauge explains: "[W]e could accept the proposition that the greatest good in life is health, but still conclude that purchasing more health care is less effective than funding nutritious food, safe housing, environmental protection, college tuition, or even simply distributing cash."[96] Yet, the United States now spends more on health care than on food, housing, transportation, or anything else.[97] Such a flexible regulatory presumption against consumption of expensive medical technologies will create opportunities for other spending that is likely to enhance patient outcomes and promote social welfare.

94. *See e.g.*, Amy Finkelstein et al., *The Oregon Health Insurance Experiment: Evidence from the First Year*, 127 Q. J. ECON. 1057, 1058-62 (2012) (showing that recipients of Medicaid coverage had self-reported improvements in well-being).

95. *See e.g.*, David M. Cutler & Mark McClellan, *Is Technological Change in Medicine Worth It?*, 20 HEALTH AFFS. 11, 11 (2001) (finding that for four conditions: "heart attacks, low-birthweight infants, depression, and cataracts—the estimated benefit of technological change is much greater than the cost.").

96. Elhauge, *supra* note 1, at 1461-62.

97. Jost, *supra* note 73, at 537.