2000-10-01

Valuing rare pediatric drugs: an economics perspective

VALUING RARE PEDIATRIC DRUGS: AN ECONOMICS PERSPECTIVE

Rena M. Conti
Jonathan Gruber
Daniel A. Ollendorf
Peter J. Neumann

WORKING PAPER 27978

NBER WORKING PAPER SERIES

VALUING RARE PEDIATRIC DRUGS: AN ECONOMICS PERSPECTIVE

Rena M. Conti
Jonathan Gruber
Daniel A. Ollendorf
Peter J. Neumann

Working Paper 27978
http://www.nber.org/papers/w27978

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2020

The origin of this study is a workshop on this topic sponsored by Sarepta Therapeutics. This study was motivated by, and summarizes many of the lessons from, this workshop. We are grateful to all participants in that workshop. Gruber, Neumann and Ollendorf acknowledge financial support from Sarepta Therapeutics. Conti did not receive financial support from Sarepta Therapeutics. This study and publication was not contingent on Sarepta's approval or censorship of the manuscript. The authors thank Kao-Ping Chua for excellent comments on a previous draft. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Valuing Rare Pediatric Drugs: An Economics Perspective
Rena M. Conti, Jonathan Gruber, Daniel A. Ollendorf, and Peter J. Neumann
NBER Working Paper No. 27978
October 2020
JEL No. I18

ABSTRACT

There is a coming wave of novel genetic therapies aiming to treat rare pediatric disease. A large literature investigates the valuation of new treatments, but the valuation of treatments for rare pediatric illness raises a host of unique issues. In this paper, we review the challenges of applying both the standard economic model and standard approaches to estimating cost-effectiveness using the quality-adjusted life year (QALY) to this case. We argue that there are a large number of special issues that have only been partially addressed by past work and we conclude that more data and the development of new methods are vital as innovators, health technology assessment practitioners and policymakers confront the launch of these new drugs.

Rena M. Conti
Boston University
Questrom School of Business
Department of Markets, Public Policy and Law
595 Commonwealth Ave
Boston, MA 02215
rconti@bu.edu

Jonathan Gruber
Department of Economics, E52-434
MIT
50 Memorial Drive
Cambridge, MA 02142
and NBER
gruberj@mit.edu

Daniel A. Ollendorf
Center for the Evaluation of
Value and Risk in Health
Tufts Medical Center
800 Washington Street, Box 63
Boston, MA 02111
dollendorf@tuftsmedicalcenter.org

Peter J. Neumann
Center for the Evaulation for
Value and Risk in Health
Tufts Medical Center
800 Washington St, Box 063
Boston, MA 02111
pneumann@tuftsmedicalcenter.org

One of the most significant developments in biology over the past 50 years is the understanding of the genetic basis for disease. We now understand that a wide variety of illness has its basis in genetic mutations. And the sequencing of the human genome in the early 2000s has led to a wave of related discoveries for the genetic basis of disease. This discovery has led to the possibility of a new type of treatment for diseases long considered untreatable through genetic therapies.

The idea of modifying DNA to treat disease is representative of a major breakthrough in medicine. Rather than treating symptoms, gene therapy aims to modify the underlying genetic etiology of a disease. The first commercially available genetic therapies, chimeric antigen receptor-T cell (CAR-T) therapies and voretigene neparvovec (Luxturna®, Spark Therapeutics), were approved by the FDA in 2017. At this point, there are less than five available genetic therapies, but there are hundreds more in development.[i]

Gene therapy treatments, however, don't come for free. A combination of forces may lead to enormous prices per unit sold of gene-based therapies.

The process of drug development is costly, risky and highly uncertain due to the trial-and-error nature of research. Every scientific discovery that becomes a new drug is the result of dozens of steps, from identifying the "target" biological entity (e.g., gene or protein), to finding the right "hit molecule" that interacts with the target and is the building block for a drug that could combat the disease, to three stages of clinical trials with the possibility of failure at each stage. To give a sense of the work involved, once a target has been identified, anywhere from two hundred thousand to more than one million compounds must be screened to identify hits, and only 9.6 percent of the drugs that enter human trials in stage one end up being approved. [ii]

Added on top of this is the much higher manufacturing costs for cellular and genetic therapies. Manufacturing cell and gene therapies is much more involved than for other pharmaceuticals. The fixed costs of setting up manufacturing capacity and the incremental costs of making each dose are both high. This is different than the costs associated with small molecule tablets and pills where the incremental costs of production are virtually zero. Moreover, whereas a single manufacturing process or platform could be deployed to produce a wide variety of traditional small molecule pharmaceuticals to treat different diseases, cell and gene therapy manufacturing is highly specialized. Not all manufacturing facilities can support the full range of manufacturing processes used, and in general, manufacturing plants for traditional drugs cannot be easily converted into plants that produce cell and gene therapy products.[iii] The cost of setting up a large-scale cell or gene therapy manufacturing facility can be

upwards of $200 million, compared with less than $1 million for a similar-scale small-molecule manufacturing facility.[iv] The quality control involved in cell and gene therapy manufacturing – making sure each and every dose provides that same safety, efficacy and effectiveness - is very intense, involves highly specialized labor and capital, and differs substantially from that entailed in traditional small molecules. Furthermore, a single facility may not be enough. Cell and gene therapy treatments are highly unstable and have extremely short shelf lives (for example, the drug Sipuleucel-T (Dendreon) has a shelf life of eighteen hours), so decentralized manufacturing at multiple smaller and geographically separate nodes (rather than a single centralized hub) becomes a necessity to preserve treatment availability and assure fidelity to high quality standards.[v]

This is all made much more complicated by the targeted nature of many of these therapies, which can be focused on very small populations facing deadly diseases. "Orphan drug" innovators must spread the costs, risks and uncertainty of their development and manufacturing across very small populations to yield sufficient financial return to motivate investment.

A final complication is that there is no natural market-based solution for the pricing new genetic therapies. To date, each therapy is unique and typically faces little competition. Consumers purchasing these therapies are often desperate and face imperfect information in assessing the underlying value of the treatment. In contrast to the U.S.'s highly successful small molecule prescription drug market, we do not expect brand-brand or brand-biosimilar competition to bring down the prices of these products.

Combined, these features may motivate manufacturers to charge very high prices. Nevertheless, pricing leverage by the innovators of these products is not unlimited. Rather, a natural benchmark for pricing is the *value* of the new therapy. Loosely speaking, the value of a new drug is the benefits to society of having this drug relative to the next best treatment. In theory, this is a readily measureable concept. The value of a drug is the incremental improvement (i.e., over the next best alternative) it delivers to the quantity and quality of life for potential patients and the associated benefits to others.

In practice, the measurement of value of a new therapy faces a number of difficult challenges. Some of these are technical, such as assessing the probabilities of various health improvements over different time frames from the treatment. Some of these involve difficult measurement concepts, such as how to properly account for the value to others in society from improving the health of an individual. And some of these involve difficult social

judgments, such as how to value a drug that creates a small life improvement for many individuals versus one that creates a large life improvement for a small number of individuals.

This paper considers how to measure value as an input to setting prices for rare pediatric disease. While most of the issues laid out here apply equally to rare diseases at any age, rare pediatric diseases raise a host of particular issues such as: do innovations for pediatric populations deserve special consideration because children represent a "vulnerable" group unable to fully speak for themselves (a rationale typically used by health assessment bodies to make exceptions for valuing pediatric drugs)? Do we take the perspective of the child or their parents? How do we deal with the fact that we could be delivering decades of extra life? How do we measure the extra caregiver burden relative to normal parenting demands? Thus, the purpose of this paper is to lay out these issues applied to the specific case of new therapies to treat rare pediatric disease and to draw some conclusions regarding how the fields of health economics and decision science may move forward in addressing them.

Part I lays out the reasons that the competitive market is unable to provide a proper signal as to the underlying value of rare pediatric drugs. In Part II, the valuation framework that is typically used as an alternative is discussed, along with common concerns and criticisms. In this section, we focus on the particular issues that arise in treating pediatric diseases. Part III then turns to the critical question of whether a higher threshold should be applied for rare diseases. Part IV applies a particular example, valuing treatments for Spinal Muscular Atrophy (SMA) that highlights the difficulties that such exercises face. In Part V, we articulate a research agenda that should be pursued in the face of the coming wave of treatments for rare disease.

Part I: Failures in the Market for Treatments for Rare Pediatric Illness

The starting point for economic modeling is the perfectly competitive market with rational, optimizing consumers. Consumers maximize a well-defined utility function that allows them to rationally trade off the costs and benefits of different baskets of goods to make themselves as well off as possible given their financial constraints. Firms maximize their profits subject to the limits of market competition, whereby perfect entry and exit serve to minimize the ability to charge prices much above the cost of producing goods.

In such a market setting, goods are purchased by consumers so long as their price is below the value of consuming the next unit of the good, and they are sold by producers so long as the price is above the cost of producing the next unit of the good. In equilibrium, the price that meets both of these conditions is the price that is

equal to the value of that next unit. The beauty of the market is that this price satisfies both consumers and producers, and as a result, the good is allocated to those individuals who value it above its price. Thus, the price paid for a good represents its value to society.

In such a framework, uncertainty does not pose a fundamental challenge. Once uncertainty is introduced, consumers will maximize expected utility, which weights the possible well-being of each outcome by the probability of that outcome. Firms likewise will evaluate their production choices given the probabilities of each production outcome.

It is well recognized in economics, however, that the market mechanism fails as a tool of pricing and allocation when there are *market failures*, or deviations from the extreme set of assumptions laid out above. As laid out by Kenneth Arrow in his classic 1963 article, the health care sector is rife with such market failures. Here we point out that the problems Arrow identified are amplified in the context of rare pediatric drugs.

Imperfect Information and the Agency Problem

The most fundamental problem facing all health care markets is *imperfect information*. The classic economics model assumes that consumers are fully informed about the attributes of a product and, based on that information, can form a proper internal valuation of the good. This is clearly not true in general in health care. The decisions facing individuals are very complicated and the tradeoffs between different treatment options are difficult to understand. Neither the efficacy of different treatments, their side effects, nor their financial burdens are readily apparent to the consumer of health care in many contexts; even if they were, the urgency often associated with health care decision-making may preclude obtaining additional information.

As a result, individuals typically rely on medical experts to make many of their health care decisions. But this reliance raises a new difficulty, the *agency problem*. The experts making these decisions consider both the patient's interest and their own interest in making recommendations. The provider interest could be influenced both by financial incentives (e.g. if different treatments lead to different levels of provider compensation) and by other professional considerations (e.g. risk of malpractice lawsuits). These agency problems are mitigated for genetic diseases given that the appropriateness of treatment is in most cases confirmed by a genetic test.

There is also another unique agency problem facing decisions over rare pediatric diseases: the decision maker (the parent) is not the party that derives most of the benefits from health improvement (the child). Parents generally consider their children's utility in making their decisions, but they may weight their own well-being too

highly – or not highly enough. A number of studies reviewed in Gruber (2016) show that parents do not appear to maximize a "family utility function" which would lead to equal weight on the well-being of themselves and their children. This can be a major consideration when there are significant caregiver burdens that are offset by treatment of rare pediatric disease. These issues get even more complicated in situations where there are multiple parent decision makers who may not have the same preferences, perhaps due to an unequal distribution of caregiver burden.

Optimization Failures

Even with perfect information, the standard model assumes that a rational forward looking consumer optimizes their consumption given the uncertainty facing different consumption opportunities. But the past few decades have seen the rise of a field of behavioral economics which questions those assumptions along a number of dimensions. This field has identified a host of "choice inconsistencies" which suggests that individuals do not make decisions as predicted by the standard economics model (Kahneman, 2011). A number of these are highly relevant to the case of rare pediatric drugs.

For example, in the standard model, individuals are assumed to make proper tradeoffs between present and future costs and benefits. If an action has a large current benefit, but an even larger future cost (when properly discounted), then it will not be undertaken. But decades of laboratory research and a growing body of economic studies demonstrate that individuals are "present biased" and do not properly weight decisions with consequences in different time periods (e.g. O'Donoghue and Rabin, 1999). Such a problem is dramatically heightened in the case of pediatric rare illness. It is difficult not to be exceptionally biased towards immediate outcomes when a child's health or life is at stake - and as a result there will be a tendency to deprioritize any long run consequences for health or financial well-being.

As another example, individuals have problems assessing frequent versus rare risks. Individuals tend to make decisions as if they think rare risks are much more likely than they are, relative to common risks (e.g. Rabin, 2000). In the context of rare pediatric diseases, that means that individuals may overvalue "home run" consequences relative to more certain but less dramatic improvements in health.

Imperfect Competition

On the supply side of the equation we have another important market failure: imperfect competition. Firms producing consumer products in a competitive market are forced to set prices at or near their costs of production,

because if they do not some other firm will undercut their price. But when firms are developing unique products – where there is very limited threat of competition - there is no such downward competitive force on the price.

This problem is particularly acute for curative treatments for rare diseases. The market power value of a patent depends on how close are the substitute treatments. Even with a patent, a treatment that has close substitutes cannot set its price much higher than its costs of production. But with truly unique genetic therapies there are limited comparable substitutes in practice, as well as inertia for firms to invest in creating a competitor given the small market size. The treatment of rare diseases has all of the features of a "natural monopoly" whereby markets naturally settle into a situation of one provider with little or no alternatives.

Liquidity Constraints and Equity Concerns

Finally, we have the problem of financing purchases. Individuals face imperfect capital markets from which they are unable to borrow the total cost of treatment. Even if they value the treatments at above their price, they can't come up with the money to pay for the treatments. As a result, most individuals are insured against risks such as the cost of treating rare diseases. But the distribution of insurance coverage is unequal. Almost 30 million Americans are uninsured and cannot finance expensive treatments, including more than four million children. Other Americans face widely varying insurance coverage rules that can make the costs of treatments, even if insured, quite large relative to income; details on coverage rules for specialty drugs are covered in Chambers et al. (2018). Millions more are insured by state Medicaid programs that are financed out of inflexible state budgets, leading to rationing of care.

This relates to concerns over equity. As a result of differences in insurance coverage, there will be enormous disparities across economic groups in their ability to finance expensive treatments for rare diseases. This is particularly concerning in the case of pediatric illness, since the child whose life is impacted by the treatment is receiving or not receiving treatment based on financial decisions made by their parents and the vagaries of coverage decisions by insurers.

Part II: A Valuation Framework for New Treatments

Given the multiple failures that make it difficult to rely on the private market mechanism to deliver proper pricing for rare pediatric diseases, we need to turn to a different mechanism. In the economics of public finance, the mechanism that is used when the market does not deliver the proper pricing signal is *cost-benefit analysis* or CBA

(Gruber, 2016). CBA is commonly applied to public sector decision making around the world. Broadly speaking, it involves measuring the benefits of an intervention relative to its costs.

When public sector decision makers are weighing alternative approaches to addressing the same unmet need, for example different alternatives for building a bridge across a span of water or providing public safety to a group of individuals, a sub-type of CBA called *cost-effectiveness analysis* (CEA) is used. The underlying analytics of *cost-benefit* and *cost-effectiveness* analyses are the same – both involve comparing benefits to costs. The difference is simply whether the end goal is to pursue all activities with a positive net benefit (benefits – costs) measured in dollar terms (*cost benefit analysis*), or whether the end goal is to evaluate a host of alternatives to accomplish a similar goal measured in the same units of non-dollar benefits (e.g. longevity, morbidity, quality of life, environmental impacts, etc.), and then pick the one with the largest positive net benefit (*cost effectiveness analysis*).

The use of cost-effectiveness in health and medicine is reviewed at length in Neumann et al. (2017). Here we will highlight the key issues it raises and the key controversies over CEA as they relate to rare pediatric illness.

The starting point for CEA, and the most difficult aspect of the exercise, is to assign a numerical value to improvements in health. This is done in two steps. The first is to assign a value to an additional year of life lived in good health. The second is to assign a value to the health state of the individual during that year, relative to good health.

Value of Life

The first step relies on decades of literature of measuring the value of a statistical life (VSL). There are a variety of approaches to estimating this value.

One approach is to use a "revealed preference" approach of examining the tradeoff between job risk and wages, or between the odds of a good saving lives versus the price paid for that good. It can also be measured from surveys that ask individuals to value reductions in the risk of death. To fix ideas, the U.S. Department of Health and Human Services, which is responsible for evaluating health-related policies in the U.S., recommended the use of a VSL of $9.3 million in 2014.[vi] To consider policies which extend life, this is converted to a life year by dividing by the typical remaining life of a typical worker, since the central estimates used to inform VSL come from job risk

tradeoff studies. So if a typical worker will live for 40 more years, this suggest a value of an additional year of life of roughly $230,000.

A huge literature on notes a wide variety of problems with measuring VSL; see Viscusi and Aldy (2003) or Gruber (2016). Most notable is the fact that the use of revealed preference approaches assumes a level of rationality in making these tradeoffs that is at odds with the literature on consumer decision making problems reviewed above. Moreover, the job-risk tradeoff estimates apply not to the typical worker, but to the worker who is deciding on whether or not to take a risky job; such workers may have tastes very different from the typical person.

Other approaches use a theoretical basis to derive the value of additional life. For example, the classic analysis of Garber and Phelps (1997) suggest a value of life-year of 1-2 times annual income, while a recent update from Phelps (2019) suggests as well 2 times annual income. Mean annual income per capita in the U.S. in 2017 was $48,150,[vii] suggesting a value of life year just below $100,000.

A major issue in applying VSL or other estimates to children, of course, is that the estimates all come from decisions made by adults, typically workers. Applying them to an additional year of life in childhood may be problematic – especially as those extra years come at younger and younger ages. It is unclear how a trade-off between job risk and wages for a 40 year-old is relevant to the value of an extra year of life at 1 year old. In particular, valuing life for small children requires directly tackling the thorny issue of considering self-reported value of life versus that reported by others (such as parents).

Quality Adjusted Life-Year

Medical interventions typically do not typically operate simply to bring someone from a 'dead' health state to good health. Rather, they often improve health among those who would have been alive but in diminished health, and they may not bring those individuals all the way to optimal health. As such, valuing the health improvements due to medical interventions requires adjusting the value of a life-year for the quality improvement during that life year. The goal of a quality adjusted life-year (QALY) framework is to bring a comparable numerical value to different states of health. The extreme benchmarks are a value of zero for dead (although some might argue that there are states of well being below being dead), and a value of one for good health. The challenge that is addressed by the QALY framework is assessing the valuation of disability and other health states that are valued less than

perfect health. Having defined the value of various health states, the analyst can compute the total QALYs gained from an intervention as a measure of the benefit to the individual.

As emphasized by Nord et al. (2009), QALYs are meant to express the personal utility of health outcomes as judged "on average" by the general public from behind a veil of ignorance about future health (so-called "decision utility"). Standard QALYs thus typically express value in terms of *ex ante* self-interest.

In practice, QALY measurement typically proceeds by using survey data to measure the relative value of health states, as summarized by Drummond et al. (2015) and Feeney et al. (Neumann et al, 2017, Chapter 7). A variety of survey measurement tools and techniques are used to value QALYs, ranging from reported rating of well-being in different states to "standard gamble" techniques where individuals are asked how much they would pay for probabilities of avoiding certain outcomes.

Implementing the QALY framework involves a very difficult set of issues that have been addressed in an extensive literature. For example, one key difficulty with the QALY approach is that alternative survey tools, and alternative ways of framing the valuation, can yield varying results for the value attached to different health states (Lipscomb, 2009); this may not be a significant issue when comparing treatments if the same survey tool is used for both, but it still can matter if the different treatments operate along different parts of the health distribution. This set of issues is covered in depth, and recommendations made, in Neumann et al. (2017) which brings together a host of experts in the area to critically review the issues.

Another difficult issue arises in using patient preferences (from directly surveying patients and their families) versus using community preferences (from surveys of the community on the value of health states) (Feeny et al., Neumann et al. 2017, Chapter 7). The latter has the advantage of reflecting social views on the value of health improvement, and avoids the very small sample sizes as well that may arise when using patients with rare diseases. On the other hand, given the uniqueness of pediatric rare illnesses, the specific views and lived experiences of the impacted patient groups may be critical.

Another major issue that must be addressed here is the extent to which individual stated preferences should in fact be used as a guide to cost effectiveness analysis. As discussed in the first section of this paper, researchers have documented a variety of "behavioral biases" in individual decision making. These biases will impact not only market decisions but survey responses as well – particularly when considering very complicated tradeoffs. A critical research issue is how to incorporate stated preferences in the context of such potential biases.

The set of issues raised by QALY measurement are particularly difficult for pediatric treatments. For example, the benefits of treatment attribute both to the child and their family – so whose opinions should be valued in establishing quality of life weights? Clearly, for diseases affecting very young children, it is impossible to get a valid self-reported measure of quality of life impacts. But just as clearly, when someone achieves a mature age, we would want to use their own self-reported quality of life measure. How do we handle cases in-between – how do we properly weight the views of children and their parents as children approach the age of majority?

On the other hand, the good news for measuring valuation for children is that the value of extra life years will in many cases simply dominate any changes in QALY. Many of the rare therapies being developed for children allow them to live many more years at a high quality of life, so VSL estimates or simple gains in life expectancy may be most of what is needed to judge cost-effectiveness.

Indeed, there is some move towards reporting not only QALYs but separately just years of life saved. This has the virtue of avoiding many of the difficult QALY measurement issues – in a sense, placing more weight on the factor which is measured more precisely (increased duration of life, rather than increased quality of life). It has the cost of overweighting the extension of disabled lives. But for rare pediatric treatments, which typically can cure a disease and extend healthy life by many years, years of life saved is likely to dominate.

The Net Costs of Treatment

Offsetting these benefits are the net costs of the intervention. On the medical side, this includes the cost of the treatment itself, as well as any associated medical costs – for example the costs of a hospitalization stay to administer the treatment. But it subtracts the savings from replacing the alternative treatment to which the intervention is being compared. At the same time, there may be a series of non-medical costs as well associated with treatment. For example, there is the time cost of the treatment itself.

In principle, this is the most straightforward part of the calculation, since it requires measuring an existing counterfactual. In practice, there are a number of difficult issues in this arena as well, particularly for pediatric rare illness.

The first problem is that the appropriate measure of reduced medical spending is defined relative to the next best treatment for the illness. But there can be wide variation in the alternative treatment patterns for pediatric rare

illnesses. This is particularly true since a large share of the spending on such illnesses may be palliative or accommodative, rather than curative, and thus not subject to traditional tests of medical appropriateness.

The second problem is that the proper measure is not the price of medical services, but their opportunity cost. For example, consider the case of services with large fixed costs and small marginal costs, like an MRI scan, as discussed by Basu (Neumann et al., 2017, Chapter 8) – should the savings of the foregone MRI be valued at the average cost of an MRI or the marginal cost of an MRI? As another example, the prices charged for medical services deviate widely from their marginal costs. We have already discussed the role of market power in the delivery of medical services, and recent research has documented the enormous variation in pricing for comparable treatments both across areas and across providers within an area (Cooper et al., 2018). If a patient lives in a very low cost area, or goes to a low cost provider, how should that be incorporated into the calculations?

Other Costs and Benefits: The Societal Perspective on Valuation

A major distinction that is drawn in valuing new treatments is between a health care sector perspective, described thus far, and a societal perspective (Neumann et al. 2017). Both perspectives have as their starting point the value of the improvement in health to the individual, and the costs of the treatment and associated medical costs, minus the savings in other medical services. The difference is that a societal perspective broadens this to consider factors such as the impact on others besides the patient, ranging from caregivers to society as a whole, and the total consumption costs and productivity gains of extending quality of life or years of life. Three issues illustrate the key differences between the health care and societal perspective – and have important implications for CEA in rare pediatric diseases versus other types of illness.

The first is reduced caregiver burden. If better health for an individual reduces the need for caregiver assistance, society benefits. One difficulty here is that this should measure the reduction in caregiver burden relative to the next best alternative. When considering caregiving of adult patients, the default is that the caregiver would spend little time caring for a healthy adult. But when considering care of children, the alternative is not zero care. All children get sick occasionally or have some risk of an acute medical event such as breaking a bone. The caregiver burden must be assessed relative to the (sizeable) burden on parents of healthy children. Moreover, reduced caregiver burden must be measured in dollar terms for inclusion in the value calculation. There are standard methods for valuing reduced caregiver burden in terms of time reduction. But reduced caregiver burden is not just

time – it also involves mental and physical well-being. This in principle requires its own set of QALY-like calculations to value, and quantifying these caregiver or family "spillover" effects where feasible (as emphasized by Neumann et al. 2017)

The second issue is that of including "related" versus "unrelated" medical costs. The classic example here is a life extending treatment that saves the costs of the current illness, but puts the patient at the risk of other illnesses that would not have occurred if they had died of the first illness (a problem economists and epidemiologists label "competing risks"). There is an extensive debate within the CEA literature on whether these "unrelated" medical costs should be counted, and the general consensus appears that they should be included in a societal perspective (Rappange et al., 2008). At the same time, this is a much smaller issue for children than for the elderly, since baseline medical spending is so low for young persons relative to older individuals. Therefore, as highlighted by Meltzer (1997), ignoring these costs will bias CEA in favor of the elderly over the young.

But a third related issue is that that once one takes a societal perspective, it is not only unrelated medical costs but a broader range of social costs – and benefits – that should be included. Essentially, the societal perspective suggests that the benefits include all of the additional productive labor that the person will provide to the market and all the additional costs of their consumption throughout their life (Meltzer, 1997; Basu, Neumann et al, Chapter 8). These measures can be very large for a treatment for pediatric illness that adds many years to life – and very difficult to measure since any productivity gains will have to be projected years into the future (offset somewhat by the fact that these will be discounted).

The distinction between the health care and societal perspectives, while often emphasized, is fairly artificial. It involves, for example, drawing difficult lines between "related" and "unrelated" health care costs, and between costs to self and to caregiver. Moreover, the health care perspective neither is the right one for any given party (e.g. a given insurer cares only about own costs; a given individual doesn't care about medical costs for which they don't pay), nor for society as whole. This distinction is particularly blurry in the context of pediatric drugs. When we value benefits to children, we necessarily consider the perspective of others, notably their family (Neumann et al. 2017).

That said, a societal perspective can quickly get overwhelming, and the discussion of this framework by Neumann et al. (2017) has been criticized by those such as Paulden et al. (2017). Whether a given person lives or dies can have ripple effects that spread broadly, from themselves, to their families, to their employers, to the

taxpayer, and even to their own future children. As noted by Saloman et al. (Neumann et al., 2017, Chapter 6), we must impose some "rule of reason" lest the calculations become unmanageable.

Perhaps the key element of any rule of reason is deciding how much of the social aspects of costs and benefits are incorporated into the QALY estimate. Basu (Neumann et al., 2017, Chapter 8) argues that the QALY framework does not incorporate any income effects from improved quality of living, so that it implicitly ignores both higher productivity and higher consumption. We are not aware of similar literature on the VSL. For a forward looking individual, the VSL should include any surplus they derive from additional years of life. Whether the VSL also includes the welfare of others is unclear; it depends on the degree of altruism of the individual and the extent to which they reflect their connections to others in their own work and purchase decisions.

Regardless, the VSL will clearly not include a number of externalities of additional life, those impacts on others that are not considered by the individual. For example, by working more, the individual may pay more taxes or collect more transfer payments from the government. The total cost of these flows are not a social cost, since they are just a transfer from one party to another, but there are inefficiencies (dead weight losses) associated with raising and spending revenues, and these should be included. Likewise, if a longer life means a higher odds of engaging in criminal activities, that has external costs on society.

One framework for addressing this set of complicated issues is Multi-Criteria Decision Analysis (MCDA). As emphasized by Phelps and Madhavan (2017), MCDA allows the user to place specific weights on different aspects of social valuation in order to be explicit about which elements matter in cost effectiveness evaluation. This approach does not "solve" the problem of multi-dimensionality and proper application of the rule of reason, but it does provide a coherent framework for making such choices transparent. An excellent example of applying MCDA to the evaluation of vaccines is discussed in Phelps et al. (2014).

Heterogeneity

The discussion thus far has proceeded under the assumption of (a) equal costs and benefits across individuals who have the same health impact and (b) homogenous impacts of treatment on costs and benefits across individuals. In this section, we discuss both of those issues and their implications for CEA.

The former issue is one that is discussed at length in the literature on VSL (Viscusi and Aldy 2003). Measured VSLs vary enormously by age, by gender, by income level, and by location. Applying these values

would suggest that the CEA for a given intervention would vary significantly depending on to whom the treatment was applied. Similarly, the costs to caregivers would vary significantly across individuals, based on the underlying opportunity cost of the caregivers' time. As another example, consider the societal approach to CEA and its treatment of productivity and consumption. As laid out by Basu (Neumann et al, 2017, Chapter 8), such an approach would count as a benefit the productivity gains of improved health, and as a cost the extra consumption (medical and non-medical) that is made as a result of improved health. In other words, such an approach would implicitly favor those who save more. And it is well known that savings rates rise with income in the U.S., so this would favor those with higher incomes.

The latter issue is more complicated and involves gathering data on treatment effect heterogeneity. This may be virtually impossible for very small sample diseases like pediatric rare diseases. At a minimum, treatment effect heterogeneity should only be incorporated into calculations to the extent that it is statistically meaningful.

Uncertainty

A proper CEA analysis must reflect the uncertainty with which benefits are delivered and costs are incurred. Sculpher et al. (Neumann et al., 2017, Chapter 11) review the incorporation of different types of uncertainty into CEA calculations. They make the important point that with multiple sources of uncertainty around key parameters, the proper approach is to specify the parameter space fully and then to probabilistically draw from this distribution of possible parameter values multiple times to understand the distribution of possible outcomes of the analysis.

As noted above, the new treatments for rare diseases come with a significant degree of uncertainty. Treatments for these small populations come with expedited FDA review that can yield approval when there is a greater than usual uncertainty about ultimate efficacy of the treatment. Moreover, given the necessarily limited nature of the trials relative to the potential gain in lifespan for pediatric rare diseases uncertainty increases as greater extrapolations of duration of benefit are required.

Discounting

Discounting is another critical issue in properly valuing treatments, particularly those for children for whom the benefits may be quite long lived. In standard economic analysis, money delivered in the future is worth

less than money delivered today, since if that money was made available today it could be invested and yield a higher amount tomorrow. That is accounted for by *discounting* future payments to recognize their relatively lower value. The major issue here is the proper discount rate to be applied to the calculation, as reviewed in Basu and Ganiats (Neumann et al, 2017, Chapter 10). CEA analysis typically use a 3% discount rate, but this is a point of some contention (Pauldsen, 2017; Neumann, et al. 2018).

Discounting in CEA has one confusing aspect. Consider two interventions:

- A: treatment cost of $100,000 that saves 1 year of life immediately
- B: treatment cost of $100,000 that saves 1 year of life in ten years

An important aspect of discounting is that the cost effectiveness of treatment B is much lower than that of treatment A. Critically, this is NOT because we value life less in the future. But rather because if we can save a year of life for $100,000 under treatment A, we could do that same thing in ten years. So instead of doing treatment B, we could take the $100,000 today, invest it, and have enough money in ten years to save more than 1 year of life.

Another important issue is the fact that the younger the child treated, the more distant into the future may go the benefits. This highlights the critical issue of discounting. In addition, there is the issue that the benefits of treatment are measured relative to the next best alternative. Given the infrequency of some pediatric illnesses, and the rapid technological change in medicine, the next best alternative may not be well defined.

Insurance Value

An important aspect of the value of new technologies that is emphasized by Lakdawalla et al, (2017) is their insurance value. As the authors point out, when a new technology is introduced, its value does not derive solely to those who get treated – but also to those who now have a higher utility from knowing that, should they have a bad health outcome, they will be cured. Insurance value considerations are standard in economic analyses of social insurance programs (e.g. Finkelstein and McKnight, 2008; Engelhardt and Gruber, 2011), but Lakdawalla point out that the same logic extends to new technologies as well.

There are two components of insurance value. The first is the physical benefit from knowing that you will be cured if becoming ill. For example, for a parent who is pregnant, knowing that a potential rare illness in their child can be cured is incredibly valuable. This benefit applies to all patients, insured or not.

There is also a second purely financial benefit: insurance is more valuable if there are expensive treatments that are covered by that insurance. For a parent who may have an ill child, the value of insuring that child is much larger when that insurance will cover treatment for a child's illness. Of course, this is only true to the extent that the new treatment is more expensive than the existing treatment; if a cure will lower total net costs of caring for a patient, then the financial benefit is actually negative. Moreover, this financial benefit will be opposite signed for an uninsured patient: an expensive new treatment increases financial risk for someone who doesn't have insurance.

Lakdawalla et etl. (2017) compute examples that suggest that these insurance values are large. They argue that the physical benefit must exceed the financial cost for an uninsured patient, but this is subject to the particular specification of their model. What is clear is that the financial insurance aspects add incremental value to insurance – insuring the population against rare illness raises the value of discovering cures for that illness.

Questioning QALYs

Some question the utilitarian perspective of just adding up QALY improvements across individuals. This will not be appropriate, for example, if there are distributional or equity considerations. This will also not be appropriate if individuals value large changes in health improvements to one person (e.g. an extra 10 years of life) differently than small changes in health improvements to many persons (one extra month of life for 120 persons).

Perhaps the most powerful broad critique of the QALY framework comes from Nobel Prize winning economist Daniel Kahneman, who rejects the *ex ante* nature of impersonal utility measurement in favor of the *ex post* concept of "experienced" personal utility. Kahneman refers to his own research which finds extreme "adaptability" among disabled persons. While *ex ante* individuals would place a huge negative value on becoming disabled, disabled persons over time in fact report a much smaller decline in well-being. This is because they are able to adapt over time in a way that cannot be foreseen by the outside observer from an *ex ante* perspective. Kahneman argues that only by measuring the actual experienced utility of different health states can we properly measure them.

This perspective is much more challenging when it comes to rare pediatric diseases. First of all, relying on experienced utility means gathering data only from those patients with the illness, and samples are small. Second, for genetic diseases there is no available baseline of "pre-disability" to which utility can be compared. Even if there

are some disability free years before illness manifests, given changes in well-being over time as children age it is impossible to make a useful comparison.

Part III: Should we Favor Treatments for Rare Diseases?

One important issue which has been raised in this area is whether we should offer different standards for rare diseases, such as a higher QALY threshold for cost-effectiveness. This is discussed at length in Jena and Lakdawalla (2017), Lakdawallala et al. (2018), Ollendorf, et al. (2018) and Garrison et al. (2019). We summarize these arguments in this section and then provide critiques of their approaches.

As a starting point, the standard economic framework would suggest that the same standard be applied to all diseases. The goal of this framework is simply to determine the value of per individual treated, or per quality-adjusted life year added. When valuing treatment, the same framework should be used regardless of incidence; otherwise you are implicitly valuing QALY improvements due to one disease more than another. This issue becomes increasingly important as treatments become more specialized. If we favor treatments with small markets, this will bias us towards more approval over time – and will bias drug development towards specialized drugs rather than general purpose ones.

There are four counterarguments made for such differential standards:

Affordability

Much discussion of the QALY framework in recent years has focused on affordability concerns. A cure for a disease that impacts a much larger population implies a much larger budgetary cost for private and public insurers. To consider a relevant recent example, consider Hepatitis C, a chronic illness impacting about 3.5 million Americans. Starting in 2013, a series of treatments which could cure this illness were introduced. The initial list price of the first of these treatments, Sovaldi, was set at $84,000. This is well below typical estimates of the value of curing Hep C.[viii] But the budgetary implications were enormous – this implied that curing Hep C nationwide would cost almost $300 billion, or about one-third of the total amount spent on health care in this country!

Compare this to the potential cure for a disease that impacts only 1000 Americans currently, such as rare pediatric diseases. Even at a cost of $1 million per treatment, this would have a much smaller budgetary impact amounting to $1 billion.

Novelty and Market Thickness

A second and related argument relates to the novelty of drugs and the "thickness" of the market for drug development. The costs of developing a new drug are largely independent of the size of the market, but the returns to developing that drug depend on the size of the market. As a result, we may have too little development of drugs for thin markets; indeed, this was the very motivation for the Orphan Drug Act.

Moreover, drug development can have spillovers – new lessons learned for treating other disease. Suppose that the spillovers arrive as constant random events independent of the size of the market for which the drug is being developed. In that case, then the fact that R&D is biased against small markets will reduce the potential positive spillovers from drug development. Indeed, the ratio of spillovers to own benefits is likely to be larger the rarer is the illness.

The Fair Innings Argument

The third argument is that the linearity of QALY measurement leads to an undervaluation of treating rare disease. As Ollendorf et al. (2018) argue, "First, some have argued that the goal of a health system, or of a society more broadly, is not simply to maximize health gains across the entire population. In this view, fairness can be defined as ensuring that all patients get some chance at a meaningful health gain (e.g., surviving a universally fatal child- hood disease), even if this exceeds standards for what would be considered a cost-effective use of health resources." This "fair innings hypothesis" can essentially be interpreted as saying that there is a non-linearity when we combine QALYs – this would say that taking one child from 0 to 1 QALY is worth much more than taking ten adults from 0.8 to 0.9 QALY. Indeed, surveys of individual preferences suggest enormous value placed on the fair innings hypothesis (Jena and Lakdawalla, 2017).

A related point is that while a disease may be rare on a population basis, it may be much less rare within a family due to genetic similarities. We may place extra value on curing a disease that can have such a devastating impact on multiple children in a family.

Insurance Value

Jena and Lakdawalla (2017) argue that one reason for social preferences for "fair innings" may be the insurance value argument developed by Lakdawalla et al. (2017). Under standard specifications of economic preferences, the value of insurance rises with both the rarity and the severity of the financial shock. Therefore, individuals may value long shot cures more than QALY-equivalent common cures because the former resolves much more physical and financial risk

Evaluating the Arguments

It is useful to distinguish the first pair of these arguments from the second pair. The first two arguments are what economists call "second best" arguments. If the goal is to preserve an efficient allocation of society's resources, both of these concerns should be addressed with other tools of social policy, not through changing the standards of cost-effectiveness analysis. For the affordability concern, efficiency dictates setting the budget so that we can treat any disease that meets cost effectiveness thresholds, rather than setting disease-specific standards. And issues such as underinvestment in novel treatments can be addressed through public R&D policy, for example for larger public subsidies (including government investments in basic research as well as development efforts and the use of prizes) for treating rare diseases (Gruber and Johnson, 2019). But in fact these alternatives may be not be politically feasible. For example, budgets are limited due to political economy constraints. And if there is no way to ensure that public R&D policy will provide proper incentives for R&D on rare diseases, then a higher cost-effectiveness threshold may be the only way to ensure that such drugs get developed.

The third argument is fundamentally different in that it comes to the basic notion of the QALY framework. Once again, the best way to address this concern would be to stick with a QALY framework but to incorporate proper weighting across disease risks. A rational way to do this would be to rely on the fourth argument – to build in insurance value into QALY calculations that incorporates "fair innings" through the higher insurance value provided by curing rare diseases. A recent paper by Lakdawalla and Phelps (2019) offers an approach to doing so through incorporating higher order preferences towards uncertain outcomes. However, this may not be enough to fully reflect social preferences.

Ultimately, the question comes down to whether the QALY framework is a fully normative perspective on how new treatments *should* be valued, or whether this is simply a measurement exercise for representing stated preferences. If the latter, then the stated social preferences for valuing rare cures more highly should be incorporated into cost effectiveness thresholds. But if the former, then these stated preferences may reflect behavioral biases which policy makers don't want to incorporate into creating drug valuations. One natural resolution to this dilemma would be to assess whether proper incorporation of insurance value is enough to approximate social preferences for fair innings.

Part III: Problems in Practice: ICER Review of SMA

Computing QALYs is hard. To see how these valuation difficulties play out in practice, consider the report by the Institute for Clinical and Economic Review (ICER) on two potential treatments for Spinal Muscular Atrophy (SMA) (ICER, 2018). ICER is a widely respected and cited source of cost effectiveness calculations for new pharmaceutical treatments in the U.S. In many respects its work parallels that done by government-sanctioned organizations in other nations, such as the UK National Institute for Health and Care Excellence (NICE), but in the U.S. ICER is a private non-profit organization. Traditionally, ICER has used a range of $50,000 to $150,000 per value QALY gained. A recent report for the UK Centre for Health Economics offers a discussion of the range as well.[ix]

SMA is rare genetic neuromuscular disorder that can be identified by mutations in the SMN gene. There are roughly 500 new cases per year, or one per 10,000 live births. The most severe cases impact infants and young children, causing muscle weakness that lead to immobility, inability to ventilate, and death. The life expectancy for the most severe type of SMA (Type I) is less than two years.

In December, 2018 ICER released a report on two potential treatments for SMA, Spinraza and Zolgensma (a revised report was released in March, 2019). Spinraza is an injectable treatment every four months which mitigates the progression of the disease, and has been approved by the FDA. Zolgensma is a new genetic treatment being developed as a one-time intravenous administration to cure SMA, and was recently approved by the FDA.

The ICER report is state of the art. In 190 dense pages it lays out the framework for applying a valuation approach to these treatments. There is an objective review of all of the available clinical studies and a detailed explanation of how available evidence is incorporated into its valuation calculations. Most importantly, the report is candid on the wide variety of limitations inherent in their exercise. There are three in particular that present enormous challenges in terms of value considerations for rare pediatric diseases.

The first is the limited size of the patient population studied. Studies are few, and the population size included in these studies is small. Moreover, only a share of studies heed to the gold standard of randomized controlled trials (RCT), whereby patients are randomly allocated to the treatment relative to a control group of similar patients who do not get the treatment. Many studies do not allocate the treatment randomly, or do not have a control group, or both.

For example, the ICER report cites two RCTs of Spinraza for Phase I SMA, each with about 80 patients receiving Spinraza and 40 controls. For Zolgensma, there is just one study, with a total of 15 patients enrolled so far, and no control group.

The second is the limited scope of existing studies. Given their small size, these studies typically do not gather data on the host of relevant indicators that would be incorporated in the full state-contingent evaluation of the benefits of treatment. This significantly hamstrings the ability of ICER and others to value health states. For example, in their valuation exercise, ICER focuses on the impacts on sitting and walking, but not on head control, rolling, crawling and standing; ICER must also assume SMA type I patients who gain the ability to walk will have a similar prognosis to SMA type III patients who already can walk. These studies are also inherently limited in their ability to project long term outcomes because of the short follow up duration; as a result, ICER makes (in its base case) the assumption that patients remain in an improved health state after treatment until death. In addition, valuing progression during the trial window is hindered by limited access to sensitive patient level data from past studies.

The third is the lack of complete and updated data on valuing the health improvements that result from new treatments. For example, a key measure of the utility value of SMA Type 1 comes from a European study with 7 patients, while the utility measure used for the sitting health state comes from clinical experts, not from patients or caregivers or general population preference surveys. Perhaps more surprising, these limitations extend as well to data on costs. For example, the data on costs associated with ventilator dependent children living at home comes from a 2002 study done in the United Kingdom.

Once again, none of these comments are meant as a particular criticism of ICER. Indeed, one strength of the report is its wide variety of sensitivity checks to assess whether its conclusions hold across a wide range of assumptions and uncertainties. The ICER conclusions about cost effectiveness for these treatments are generally quite robust; for example, the report finds that Zolgensma is relatively cost effective even with assumption of $2 million up front cost. Despite the limitations faced by ICER, their recommendations have a powerful effect; the initial price of Zolgensma was set at $2.1 million, essentially the cost-effectiveness threshold set by ICER.

These issues are not unique to SMA. Other ICER reports on genetic treatments for rare diseases face similar issues. As another example, consider the ICER report on CAR-T treatments for B-Cell Cancers. This is another excellent study which includes a comprehensive review and application of the available studies. But the

studies that they have to rely on are all single arm studies without randomized controls, and the studies do not last long enough to assess remission patterns.

Part IV: Forward Looking Data Collection and Methods Development

We are at the edge of giant wave of genetic treatments for rare, and not-so-rare, diseases. Health economists and organizations like ICER are positioned to evaluate the value of these new treatments. In order to increase these evaluations' effectiveness, it would be valuable to move forward now on a comprehensive data collection effort that prepares us for the coming wave.

Many of the difficult issues raised above suggest the importance of having a large volume of data to address valuation issues. The fact that there is such variation in the medical treatment of rare pediatric diseases, for example, shows the importance of having a large sample. The issue of transitioning from parental valuation to child valuation requires sufficient data at various ages to consider both perspectives.

Three important initiatives could help here. The first is to be more forward looking in our collection on the consequences for well-being of disease and its treatment. If we were planning ahead for aggressive cures to SMA, more data could have been collected on the quality of life implications of different disease states – and we wouldn't have to rely on decades old cost data on ventilation treatment.

In particular, institutions could be built and supported in the near term that will focus on collecting the data that is needed, supported in no small part by the manufacturers that are developing new therapies and sponsoring the clinical studies. This starts with a horizon-scanning exercise on a disease-by-disease basis to understand what is known and what is not about the elements of the valuation expression. This could be followed by data collection initiatives to fill those missing pieces.

Each disease can come with a unique profile of disabilities that can interact in ways unique to that illness. Therefore, existing data on valuing disabled life years may be not be sufficiently comprehensive to measure the combinations that are needed in the case of particular rare diseases. On the other hand, given the rarity of these illnesses, there is a tradeoff between collecting data from the small number of families that have experienced this disease versus a larger sample who can hypothetically weigh in on the value of health improvement from new treatments.

Some of this would require new survey and data collection from patients. Population databases are not likely to be very helpful when dealing with specific rare illnesses. Setting a goal of collecting a minimum sample size of data (e.g. 200 patients) for each illness would avoid having much weaker data to use for evaluating treatment of rare illnesses. These samples can be readily identified by taking advantage of registers created by patient advocacy groups. Of course, for truly rare diseases, even sensible minimum sample sizes may exceed the target population worldwide. In those cases, we would have to work with smaller samples. Some would involve taking better advantage of data that already exists in the health care ecosystem – such as claims data on patient treatments – and that is now being more widely used by researchers.

Such an effort would be expensive – but small relative to the scope of the issue as it will emerge in the coming decades. Suppose that twenty years from now we will be able to treat 100 new diseases that impact on average 1000 people per year at an average cost of $500,000 per person. That is spending of $50 billion per year. At that price, getting valuation right will be worth a lot of money: collecting data that allows us to improve valuation by 10% would be worth $5 billion per year! Indeed, Sculpher et al. (Neumann et al., 2017, Chapter 11) provides a framework for quantifying the expected value of additional information than can then be compared to the cost of research.

Given the dollars at stake, it is critical that broader data be collected. But there might be concerns with such data collecting being funded by interested parties. Objective collection and analysis requires a funding stream that is independent of the companies funding specific drug trials.

One option is to follow a formulation already used by the FDA to finance its work: a user fee on drug companies. The resources involved in providing richer data collection and analysis are fairly small relative to the size of the U.S. pharma market, so a very large user fee would not be required. But this would be controversial since if the government were collecting the funds, it would need to be in charge of their dissemination, and this would start to cross the politically fraught boundary into rationing discussions.

What may be required instead is some financing mechanism that is outside of the government purview. Whatever the solution here, it is critical that it move forward as the wave of new treatments for rare diseases arrives.

Second, it is also critical to invest in the science of drug evaluation. For example, as the ICER reports emphasize, the randomized controlled trial is the gold standard for evaluating efficacy. But this is a taxing standard

for rare treatments - these RCTs are expensive to run, can be difficult to recruit for, have small samples, and are often ended quite quickly for ethical reasons.

We can potentially improve on this model through considering more adaptive evaluation of treatments. The NEWDIGS initiative at MIT among others are working on ways of incorporating data from not just trials but claims data and electronic health records to evaluate the efficacy of treatments after approval and launch.[x] In the case of SMA, ICER compared the results from a randomized trial to the results from patients newly treated under expanded access programs (EAP) (ICER, 2018).. They found fairly similar results in this context, allowing them to pool the findings for much more power in the analysis.

As we move towards more analysis of treatments for rare diseases, we should be conscious of the tradeoff between gold standard data that is quite limited and "silver standard" data that may be much more expansive. We can do considerable work now to understand this tradeoff and assess the optimal combinations of data to provide evidence on treatment efficacy.

Finally, there is the issue of information from the drug trials themselves. The ICER SMA report highlighted the limitations that they faced on getting patient progression data from the trials due to patient confidentiality concerns. At the same time, other data was received from manufacturers in confidence, but under ICER's Data-in-Confidence policy has to be released at the earlier of (a) publication or public presentation; or (b) 18 months following the posting of the report. There is once again a tradeoff between making data public and getting analysts the data that they need to make the best possible evaluation of treatment effects. More work is needed in evaluating this tradeoff.

References

Arrow, Kenneth (1963). "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53 (941-973).

Chambers, James, David Kim, Elle Pope, Jennifer Graff, Colby Wilkinson and Peter Neumann (2018). "Specialty Drug Coverage Varies Across Commercial Health Plans in the U.S.," *Health Affairs*, 37, 1041-1047.

Cooper, Zack, Stuart Graig, Martin Gaynor, and John Van Reenen (2018). "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured," *Quarterly Journal of Economics*, 134, 51-107.

Garrison, Louis, Tristen Jackson, Douglas Paul, and Mike Kenton (2019). "Value-Based Pricing for Emerging Gene Therapies: The Economic Case for a Higher Cost-Effectiveness Threshold," *Journal of Managed Care and Specialty Pharmacy*, published online February 19 , 2019.

Gruber, Jonathan (2016). *Public Finance and Public Policy, 5th Edition*. New York: Worth Publishers.

Gruber, Jonathan and Simon Johnson (2019). *Jump-Starting America: How Breakthrough Science Can Revive Economic Growth and the American Dream.* New York: Public Affairs.

Institute for Clinical and Economic Review (2018). *Spinraza and Zolgensma for Spinal Muscular Atrophy: Effectiveness and Value*, December 20, 2018 (updated March, 2019).

Jena, Anupam and Darius Lakdawalla (2017). "Value Frameworks for Rare Diseases: Should They Be Different?," *Health Affairs Blog*, April 12, 2017.

Kahneman, Daniel (2011). *Thinking Fast and Slow.* New York: Farrar, Straus and Giroux

Lakdawalla, Darius, Anup Malani and Julian Reif (2017). "The Insurance Value of Medical Innovation," *Journal of Public Economics*, 145, 94-102.

Lakdawalla, Darius, Jalpa Doshi, Louis Garrison, Charles Phelps, Aniran Basu, and Patricia Danszon (2018). "Defining Elements of Value in Health Care – A Health Economics Approach: An ISPOR Special Task Force Report," *Value in Health*, 21, 131-139.

Meltzer, David (1997). "Accounting for Future Costs in Medical Cost-Effectiveness Analysis," *Journal of Health Economics*, 16, 33-64.

Neumann, Peter, Gillian Sanders, Louise Russell, Joanna Siegel, and Theodore Galniats (2017). *Cost-Effectiveness in Health and Medicine, Second Edition.* New York, NY: Oxford University Press.

Nord, Erik, Normal Daniels and Mark Kamlet (2009). "QALYs: Some Challenges," *Value in Health*, 12, S10-S15.

O'Donoghue, Ted, and Matthew Rabin (1999). "Doing it Now or Later," *American Economic Review*, 89, 103-124.

Ollendorf, Daniel, Richard Champan and Steven Pearson (2018). "Evaluating and Valuing Drugs for Rare Conditions: No Easy Answers," *Value in Health*, 21, 547-552.

Paulden M, O'Mahony JF, McCabe C. (2017). "Discounting the Recommendations of the Second Panel on Cost-Effectiveness in Health and Medicine," *Pharmacoeconomics*, 35(1):5-13.

Phelps, Charles and Guruprasad Madhavan (2017). "Using Multicriteria Approaches to Assess the Value of Health Care," *Value in Health*, 20, 251-255.

Phelps, Charles, Gururprasad Madhavan, Kinpritma Sangha, Rino Rappuouli, Rita Colwell, Rose Marie Martinez, Patrick Kelley and Lonnie King (2014). "A Priority-Setting Aid for New Vaccine Candidates," *PNAS*, 111(9), 3199-3200.

Quinn, Casey, Colin Young, Jonathan Thomas, Mark Trusheim, MIT NEWDIGS FoCUS Writing Group (2019). "Estimating the Clinical Pipeline of Cell and Gene Therapies and Their Potential Impact of the U.S. Healthcare System," *Value Health*, 22, 621-626.

Rabin, Matthew (2000). "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 68, 1281-1292.

Rappange, David, Pieter HM van Baal, N. Job A van Exel, Talitha L. Feenstra, Frans F.H. Rutten and Werner B. F. Brouwer (2008). "Unrelated Medical Costs in Life-Years Gained: Should They be Included in Economic Evaluations of Healthcare Interventions?," *Pharmcoeconomics*, 26, 815-830

Viscusi, W. Kip and Joseph Aldy (2003). "The Value of a Statistical Life: A Critical Review of Market Estimates from Around the World," *Journal of Risk and Uncertainty*, 27(1), 5-76.

[i] Quinn et al. (2019).

[ii] This section summarizes the excellent discussion of early-stage research in James P. Hughes, Stephen Rees, S. Barrett Kalindjian, and Karen L. Philpott, "Principles of Early Drug Discovery," British Journal of Pharmacology 162, no. 6 (2011), https://www.ncbi.nlm.nih.gov/pubmed/21091654.

[iii] "Manufacturing Challenges for Cell and Gene Therapies," Cell and Gene Therapy Catapult, February 21, 2017, https://ct.catapult.org.uk/news-media/general-news/manufacturing-challenges-cell-and-gene-therapies.

[iv] Ralf Otto, Alberto Santagostino, and Ulf Schrader, "Rapid Growth in Biopharma: Challenges and Opportunities," McKinsey & Company, December 2014, https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/rapid-growth-in-biopharma.

This study reports manufacturing set up costs of $200-$500 million. These costs have been falling recently, with Amgen's new state-of-the-art facility in Rhode Island expected to cost about $200 million: (https://www.amgen.com/media/news-releases/2018/07/amgen-breaks-ground-on-next-generation-biomanufacturing-plant-in-rhode-island/)

[v] Peter Olagunju, Rodney Rietze, and Dieter Hauwaerts, "Meeting the Cell Therapy Cost Challenge with Automation," Invetech, February 21, 2017, https://www.invetechgroup.com/insights/2017/02/meeting-the-cell-therapy-cost-challenge-with-automation/. PROVENGE (sipuleucel-T) is a prescription medicine that is used to treat certain patients with advanced prostate cancer.

[vi] U.S. Department of Health and Human Services. 2016. "Guidelines for Regulatory Impact Analysis." https://aspe.hhs.gov/system/files/pdf/242926/HHS_RIAGuidance.pdf. ASPE does not recommend adjusting the VSL for the age of the individual.

[vii] https://www.thebalance.com/income-per-capita-calculation-and-u-s-statistics-3305852

[viii] ICER report available at https://icer-review.org/wp-content/uploads/2016/02/CTAF_Hep_C_Apr14_final.pdf

[ix] https://www.york.ac.uk/che/research/teehta/thresholds/

[x] https://newdigs.mit.edu/programs-projects/adaptive-biomedical-innovation