

Traditional Value Assessment Methods Fail Communities of Color and Exacerbate Health Inequities

Introduction

As healthcare costs in the United States have continued to grow, there has been an increasing desire to move towards “value-based” care in which patients and payers would be paying for the “value” of the treatment patients receive. This idea has been gaining particular momentum as policymakers look for strategies to curtail spending on pharmaceuticals. While there is merit in paying for services and treatments that work and eliminating wasteful spending, it is important to step back and consider to whom “value” is being provided. As the ultimate beneficiary, we would advocate that the measure of “value” in a healthcare setting should focus on value to the patient, but currently, we are concerned that the trend is to look towards “value” to the payer by prioritizing reduced cost over care outcomes for patients in the form of cost-effectiveness analyses. For this reason, there is debate within health care stakeholders as to whether value assessments are necessary or useful. The flaws of cost-effectiveness analyses that are currently in use become particularly pronounced when evaluating treatments for population and patient cohorts that, historically, have not been included in the research, and for diseases that have not received equitable consideration in research or in diagnosis and treatment paradigms. To a significant degree, these flaws become manifest in communities and patients of color who invest in the system through taxes, insurance premiums and out-of-pocket costs, populations whose numbers are rapidly increasing the diversity of the American general population.

This issue brief will examine ways that current methods of cost-effectiveness analysis and value assessment fail to reflect true value to patients and will offer recommendation regarding methods to address this inequity.

The Quality-Adjusted Life Year and its discriminatory implications

There are many different methods for assessing the “value” of a medical treatment or intervention, but the traditional and most frequently used method relies on a cost-effectiveness analysis using the quality-adjusted life year (QALY) metric. More and more frequently, U.S. policymakers are looking to this type of assessment as a strategy to determine the true “value” of a drug and ultimately lower drug prices without weighing the fact that cost-effectiveness assessments generally do not accurately capture “value” to the patients in question, particularly for diseases with heterogenous patient populations. This method is used widely outside of the United States to determine whether, and under what circumstances, treatments and procedures should be made available and covered for patients. The most well-known example of this is the National Institute for Health and Care Excellence (NICE) in the United Kingdom. In the United States, value assessments are not conducted by the federal government, though they are used by insurer’s pharmaceutical and treatment (P&T) committees that contract with federal government programs and the Veterans Administration. A number of non-governmental agencies perform their own value assessments in the United States – the most recognizable being the Institute for Clinical and Economic Review (ICER).

Traditional methods for cost-effectiveness analyses that rely on QALYs, such as the method used by ICER, are controversial. The use of a QALY as a proxy for health gain has been widely criticized for discriminating against patients and people with disabilities, ignoring patient-centered outcomes, and failing to account for patient heterogeneity.¹ Such methodological flaws become more pronounced when trying to understand the value of a treatment for a populations not represented to a statistically significant degree in the clinical trials or the peer-reviewed literature, including, but not limited to Latinx populations, persons of African descent, and Asian-Pacific Islander and Native American populations.

¹ Ubel PA, Richardson J, Prades JL. Life-saving treatments and disabilities: Are all QALYs created equal?. International journal of technology assessment in health care. 1999 Oct;15(4):738-48.

The QALY is discriminatory towards people with disabilities and chronic conditions because it uses a weighted health utility that at baseline tends to be lower when assigned to patients who are considered to have worse health.² This has been highlighted as affecting numerous patient cohorts; from the disabled,³ seniors,⁴ and those with mental illness.⁵ The inevitable outcome is that the average QALY gains for interventions targeted at these patients will systematically produce lower estimates of overall health gain compared to patients who do not have the same health challenges. This problem is so widely accepted that there has been longstanding objection by policymakers and the American public to the use of QALY-based cost-effectiveness analyses. Notably, in 1992, the administration under President George H.W. Bush established that Oregon's efforts to use a QALY-based cost-effectiveness standard in Medicaid would violate the Americans with Disabilities Act (ADA) by systematically disadvantaging people with pre-existing disabilities.⁶ The Affordable Care Act (ACA) also includes a safeguard that prohibits Medicare from using cost-effectiveness assessments to make any coverage or payment decisions.⁷

Even those who conduct cost effectiveness analyses have recognized these problems, and there have been a number of attempts to develop alternatives to the QALY to overcome them, such as the Equal Value of Life Years Gained (evLYG)⁸ and most recently the Health Years Total (HYT).⁹ While both are successful in reducing the impact of distributional inequity as compared to the QALY, they also have wider limitations, which make them questionable substitutes for a broken measure. The evLYG limits the value of interventions that both extend life and improve quality of life,¹⁰ whereas the HYT, while improving distributional probity, does not altogether fix it.

In this age of personalized and precision medicine, groundbreaking therapies and cures are being developed for sickle cell disease, hepatitis C, heart failure, type 2 diabetes. These are innovative approaches that advance the science and enable therapeutic responses to be responsive to populations that have historically been defined as minorities in the United States. Unfortunately, conventional value assessment methodologies and metrics are at best dismissive of these patients and, at worst, willfully blind with a callous disregard for their impact on these patients.

The inherent bias problem

Most cost-effectiveness analyses rely on data from randomized clinical trials (RCTs) and health utility preference weighting surveys. These data sources rely on inputs from Caucasian populations.

Table 1: Racial breakdown of data sets used in ICER assessments in comparison to racial breakdown of the actual United States population of need.

ICER ASSESSMENT	RACE BREAKDOWN WITHIN PUBLISHED TRIAL DATA SETS		RACE BREAKDOWN WITHIN THE WIDER POPULATION OF NEED (US)	
	African American	Caucasian	African American	Caucasian
High Cholesterol	5%	86%	12.6%	72.4%
Type 2 Diabetes	7%	79%	16%	69%
Hepatitis C	7.2%	85%	22%	64%
Prostate Cancer	10%	74%	20%	80%
High Cholesterol	3%	85%	14%	69%
Congestive Heart Failure	6%	82%	16%	66%

2 Wagstaff A. QALYs and the equity-efficiency trade-off. *Journal of health economics*. 1991 May 1;10(1):21-41.

3 Amundson R. Disability, ideology, and quality of life. *Quality of life and human difference*. Genetic Testing, Health Care and Disability. 2005 May 9:101-20.

4 Cookson R, Culyer AJ. Measuring overall population health: the use and abuse of QALYs. *Evidence-based public health: effectiveness and efficiency*. 2010;148.

5 Ashton T, Cumming J, Devlin N. *Prioritising health and disability support services: principles, processes and problems*. Wellington, New Zealand: National Health Committee. 1999 Feb.

6 Brock D. Ethical issues in the development of summary measures of population health states. In: Field MJ, Gold MR, eds. *Summarizing Population Health*. Washington, DC: National Academies Press; 1998:63-81.

7 PPACA. Sec. 6301\1182 SSA. Page 678. Available at: <http://housedocs.house.gov/energycommerce/ppacacon.pdf>. Accessed January 5, 2020.

8 Nord E, Pinto JL, Richardson J, Menzel P, Ubel P. Incorporating societal concerns for fairness in numerical valuations of health programmes. *Health Economics*. 1999;8:25-39

9 Basu A, Carlson J, Veenstra D. Health Years in Total: A New Health Objective Function for Cost-Effectiveness Analysis. *Value in Health*. 2019 Dec 24.

10 Johannesson M. Should we aggregate relative or absolute changes in QALY? *Health Econ*. 2001;10:573-577.

QALY-based value assessments, particularly those developed prior to the launch of a treatment, rely heavily on randomized clinical trials for key inputs into their calculations.¹¹ RCTs are often highly homogenous, primarily consisting of white males. Fewer than 10% of Americans are enrolled in RCTs, and lack of enrollment is even greater among patients and communities of color.

There have been numerous studies highlighting this lack of diversity in RCTs both in Europe¹² and the United States.^{13, 14}

Despite representation in clinical trials being an area of increased focus by stakeholders including the FDA and biopharmaceutical industry, representation of diverse populations remains a challenge. In fact, only 8% of clinical trial enrollees are Hispanic, when Hispanic Americans make up 17% of the overall population. This lack of representation holds especially true for certain diseases. For example, Americans of African descent are more likely to have respiratory ailments than white Americans; but, as of 2015, only 1.9% of all studies of respiratory disease included this population.

Oncology is particularly illustrative of this systemic inequity. In a review of recruitment in phase 3 cancer clinical trials¹⁵ conducted between 2001 and 2010, reviewers found that just 6.2% were African American; in contrast to studies conducted between 1990 and 2000, in which 10.5% were African American. This illustrates that despite awareness of the problem that RCTs are highly homogeneous, efforts to enroll African Americans appear to have diminished or become less effective over the last two decades.

This failure to power clinical trials to yield statistically significant results for patients of color compromises the clinical validity of data and information regarding disease presentation and therapeutic responses and findings regarding safety and efficacy. By extension, methodologies designed and employed by health economists to assess the relative value of health technologies are equally compromised and discreditable.

One can easily see how this problem with scientific validity has the result of diserving communities of color and actually increasing disparities by looking at the example of ICER's assessment for type 2 diabetes shown in table 1. The assessment relied on data sets in which African Americans were underrepresented by over 50% compared to the actual population of need. This immediately presents a problem in which these results are not representative of the actual population and will serve to exacerbate disparities for minority patients.

For example, if a particular therapy is effective for an African American population, but less effective for a Caucasian population, but the enrolled trial cohort is dominantly Caucasian, with African Americans under-represented, its average effect sizeⁱⁱ demonstrated by the RCT will be small, and the therapy will have a lower chance of being approved.

Conversely, if a therapy is highly effective for Caucasians and less effective for African Americans, with a similar distribution of RCT participants as before, this will result in an overestimation of the effect size and increase its chance of being approved. Multiply this effect by the hundreds and thousands of trials that have evaluated the thousands of therapies that have been approved – or not – over the decades and you have a systematic bias of available therapies that favor Caucasians to the detriment of African Americans and other disenfranchised patients and communities.

Value assessments are largely based on population-level averages and rarely report results specific to minorities

Value assessments are largely based on population-level averages, and rarely are results specific to patients or communities of color identified in final results. In fact, in a recent review of cost-effectiveness analyses published through 2016, only 19% reported patient subgroup results and only 4.4% reported on race or ethnicity specifically.¹⁶ As a result, based on the data inputs to these assessments, the risk profile for an individual is likely to be a proxy closely aligned with someone white, middle aged, and male.

11 Editors, The. "Clinical Trials Have Far Too Little Racial and Ethnic Diversity," Scientific American, Scientific American, 1 Sept. 2018, www.scientificamerican.com/article/clinical-trials-have-far-too-little-racial-and-ethnic-diversity/.

12 Allmark P. Should research samples reflect the diversity of the population?. *Journal of medical ethics*. 2004 Apr 1;30(2):185-9.

13 Corbie-Smith G, Thomas SB, George DM. Distrust, race, and research. *Archives of internal medicine*. 2002 Nov 25;162(21):2458-63.

14 Flaskerud JH, Nyamathi AM. Attaining gender and ethnic diversity in health intervention research: cultural responsiveness versus resource provision. *Advances in Nursing Science*. 2000 Jun 1;22(4):1-5.

15 Kwiatkowski K, Coe K, Bailar JC, Swanson GM. Inclusion of minorities and women in cancer clinical trials, a decade later: have we improved?. *Cancer*. 2013 Aug 15;119(16):2956-63.

16 Lavelle TA, Kent DM, Lundquist CM, Thorat T, Cohen JT, Wong JB, Olchanski N, Neumann PJ. Patient variability seldom assessed in cost-effectiveness studies. *Medical Decision Making* 2018;38(4):487-94

It is well established that generating and reporting differential value assessment across patient subgroups leads to substantial health gains, both through treatment selection and coverage.^{17,18} Therefore, ignoring heterogeneity in patient populations, and simply providing results for overall populations—even when clinical evidence shows differential effectiveness across subpopulations—leads to a cost effectiveness assessment of value by payers that is different than the clinical assessment of value made by clinicians and patients. This disconnect can ultimately lead to inefficient decision making based on reimbursement and pricing, which disservices minority patients by not providing them with the most clinically effective healthcare. This is unacceptable.

QALYs assign lower value to the lives of patients with disabilities and chronic conditions

It is well documented that the QALY places lower value on the life years of individuals with disabilities and chronic conditions.¹⁹ As described earlier, the QALY metric is constructed using a weighted utility that is lower for those with worse health and has a discriminatory effect on those with disabilities and chronic conditions.

African Americans have higher rates as well as earlier onset of chronic conditions and disabilities.²⁰ For example, 25% of African Americans have diagnosed disabilities, and Hispanic Americans have significantly higher rates of certain chronic conditions including diabetes and obesity.²¹ Therefore, the QALY method further exacerbates health status disparities by inherently de-valuing life years lived with a disability or chronic disease.

One recent example of this issue is in ICER's assessment of treatments for sickle cell disease (SCD), a chronic and disabling disease that disproportionately impacts African Americans and Hispanic Americans to a greater degree than Caucasian, non-Hispanics. Until very recently, there has been little progress in available therapeutic options. An overdue sign of hope came in the past two years, with three new treatments for SCD approved for use by the FDA. This was a huge step forward for patients and clinicians who feel the disease has been neglected due to systemic racism. Following this new innovation, ICER initiated a QALY-based value assessment of these treatments before they were all approved by FDA, despite the shortcomings of the metric in assessing treatments for minorities. The ICER assessment acknowledged that all three treatments evaluated resulted in higher life expectancy than usual care alone, but its QALY-based economic base case still found these therapies to not be cost-effective.

Health utilities are typically derived from homogenous Caucasian non-Hispanic populations

The calculation of a QALY relies on the use of numeric "utilities" to determine the "value" of various health states. The utilities from which health states are determined are based on large, country specific surveys of the general public. These health state valuation studies that are translated into QALYs are undertaken in predominantly white populations, and weighting calculations are largely constructed using linear regression, which can over-homogenize weights around the medium health states. That means the value of a health state and its weight in the calculation of value is derived from populations that don't reflect minorities. In the design and construction of the tool that measures these utilities, the selection and construction of the "domains" that make up quality of life tools were constructed by a small group of white men forty years ago in Switzerland.²² No one has challenged these sets to be updated, as it would be inconvenient for the method despite its need to appropriately measure disease impacts in minority populations.

Utility designs may not incorporate outcomes that matter to patients, including social determinants of health

This type of utility design, by definition, over-values the largest subpopulation in the general population, which, in the United States, has been, historically, Caucasian non-Hispanics. Populations of color generally have greater exposure to risk factors,²³ greater burden from disease,²⁴ and higher levels of severity of disease.²⁵ This is primarily because they have

17 Basu A. Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. *Journal of health economics*. 2011 May 1;30(3):549-59.

18 Espinoza MA, Manca A, Claxton K, Sculpher MJ. The value of heterogeneity for cost-effectiveness subgroup analysis: conceptual framework and application. *Medical Decision Making*. 2014 Nov;34(8):951-64.

19 Brock DW. Cost-effectiveness and disability discrimination. *Economics & Philosophy*. 2009 Mar;25(1):27-47.

20 Thorpe, Roland J Jr et al. "Relationship Between Chronic Conditions and Disability in African American Men and Women." *Journal of the National Medical Association* vol. 108,1 (2016): 90-8. doi:10.1016/j.jnma.2015.12.012

21 "Hispanic Health." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, 5 May 2015. www.cdc.gov/vitalsigns/hispanic-health/index.html.

22 Oliver A. *Discovering the QALY: Or How Rachel Rosser Changed My Life*. Alan Williams in *Personal Histories in Health Research*, ed. Adam Oliver. Nuffield Twst. London 2005.

23 Leigh JA, Alvarez M, Rodriguez CJ. Ethnic minorities and coronary heart disease: an update and future directions. *Current atherosclerosis reports*. 2016 Feb 1;18(2):9.

24 McKenna MT, Michaud CM, Murray CJ, Marks JS. Assessing the burden of disease in the United States using disability-adjusted life years. *American journal of preventive medicine*. 2005 Jun 1;28(5):415-23.

25 Crosslin KL, Wiginton KL. The impact of race and ethnicity on disease severity in systemic lupus erythematosus. *Ethnicity & disease*. 2009 Feb;19(3):301.

lower access²⁶ to—and as such lower utilization^{27,28} of—healthcare services, and the healthcare services they do receive are largely poorer in quality. The current novel coronavirus crisis exemplifies the systemic inequities that result in increased risk for avoidable morbidity and mortality for communities of color, from case finding, diagnostic testing for COVID-19, and inpatient triage plans states and hospitals have put into place which favor “healthier” patients who have the potential to be “most productive” upon recovery.^{29,30}

Traditional value assessment methodologies, which rely on largely Caucasian population averages assume levels of healthcare access and utilization, treatment delivery patterns, disease burden and severity, presence of co-morbidities, and adherence to therapies that will, almost without fail, lead to an underestimate of the value of such therapies to patients of color.

An example of this is ICER’s SCD assessment, which used claims data to determine the number of acute pain crises (APC) in SCD patients. Claims data sets have the advantage of showing real-world data for cost. Nevertheless, there are many variables for which this type of data is clearly a poor proxy. No one assumes that all episodes of disease are treated in the healthcare system, especially with varying accessibility to health care within populations. As noted above, patients and communities of color have lower rates of utilization of healthcare services. This pertains to the calculation of APC for SCD. Most patients prefer to handle APC in a home setting, not through hospital admission. The American Society of Hematology (ASH) noted this to ICER in its first comment letter, and the point is backed up empirically by studies that show patients managed most of their APC at home versus in a hospital setting.³¹

Failure to incorporate outcomes that matter to patients, including lost productivity and caregiver burden also means outcomes such as lost productivity and caregiver burden are not value assessment, but it becomes especially problematic when conducting evaluations of treatments for communities of color, as research has shown that, generally, these patients have less savings, are less likely to be in salaried professions, are less likely to get sick leave or have paid sick leave, have smaller social support networks and are less likely to be insured. Therefore, indirect effects and outcomes can impact minorities in more pronounced ways.³²

One strategy to better capture outcomes that matter to patients is using disease-specific patient-reported outcome (PRO) tools. When assessing treatments for a disease like SCD that impacts primarily communities of color, disease specific PROs would solve for some of these inequities, but cost-effectiveness analyses generally rely on generic PROs. Using a general tool typically fails to capture outcomes that actually matter to the patients in question, and we again end up in a situation where a poor proxy is being used. Generic PRO tools like EQ5D especially fail to capture unique needs of minority communities. An example of this can be seen in ICER’s SCD assessment, which relies on studies that were undertaken using a generic patient-reported outcome (PRO) tool called the EQ5D, rather than disease specific tools.

Innovative methods under development that could remedy inequities presented by value assessments

From the perspective that the primary goal of a healthcare system should be to serve patients and provide them with the best possible care, we should not continue using one-size-fits all value assessments. It is imperative that the entire healthcare ecosystem work together to develop and nurture alternative approaches that reduce risk and increase the value of health care for patients, including communities of color and those with disabilities. Our healthcare system has entered an era of personalized medicine, and our goal should be to realize the promise of this evolution and ensure that each patient is getting the best, most appropriate treatment for their individual needs. At this point, health evaluation techniques serve to exacerbate health disparities across populations through “intervention-generated inequality.”³³ In order to avoid this, we must develop evaluation techniques that account for these disparities and allow us to quantify them, and then use to them to address disparities proactively.

26 Clemans-Cope L, Kenney GM, Buetgens M, Carroll C, Blavin F. The Affordable Care Act’s coverage expansions will reduce differences in uninsurance rates by race and ethnicity. *Health Affairs*. 2012 May 1;31(5):920-30.

27 Toth PP, Stevens W, Chou JW. Why published studies of the cost-effectiveness of PCSK-9 inhibitors yielded such markedly different results. *Journal of medical economics*. 2017 Jul 3;20(7):749-51.

28 May FP, Rolston VS, Tapper EB, Lakshmanan A, Saab S, Sundaram V. The impact of race and ethnicity on mortality and healthcare utilization in alcoholic hepatitis: a cross-sectional study. *BMC gastroenterology*. 2016 Dec 1;16(1):129.

29 The Times Editorial Board (2020) ‘Editorial: Who do we save from coronavirus and who do we let die? Take wealth, race and disability out of that brutal equation’. *The Los Angeles Times*, 25 April.

30 Sitrin C, Sutton S. “Why New Jersey’s ventilator guidelines may favor younger, whiter patients.” *Politico*, April 19, 2020. <https://subscriber.politicopro.com/article/2020/04/19/why-new-jerseys-ventilator-guidelines-may-favor-younger-whiter-patients-1277682>

31 Smith WR, Penberthy LT, Bovbjerg VE, McClish DK, Roberts JD, Dahman B, Aisiku IP, Levenson JL, Roseff SD. Daily assessment of pain in adults with sickle cell disease. *Annals of internal medicine*. 2008 Jan 15;148(2):94-101.

32 Choudhury, Sharmila. “Social Security Administration.” *Social Security Administration Research, Statistics, and Policy Analysis*, Social Security Office of Policy, 1 June 2003, www.ssa.gov/policy/docs/ssb/v64n4/v64n4p1.html.

33 Lorenc T, Petticrew M, Welch V, Tugwell P. What types of interventions generate inequalities? Evidence from systematic reviews. *J Epidemiol Community Health*. 2013 Feb 1;67(2):190-3.

Experts are working to develop more nuanced, patient-centric methods to do this that would identify and incorporate concerns about health disparities instead of exacerbating disparities by ignoring them.

Several organizations have set out to develop Multicriteria Decision Analysis (MCDA) as an alternative form of value assessment,³⁴ which could allow researchers to use health disparity as a criterion to evaluate new therapies directly alongside any measure of efficiency alone.³⁵ There are many examples of MCDA being used in healthcare priority-setting exercises that have included addressing health disparities both in the United States³⁶ and around the world. Examples of this are ranking alternative public health investment priorities by the distribution of benefits accrued to the most disadvantaged,³⁷ and benchmarking health disparities across geographical areas.³⁸

Other organizations are moving value assessments away from population-based averages with the goal of ensuring representation of the specific patients impacted.³⁹ These organizations, including the Patient Centered Outcomes Research Institute (PCORI), are conducting large scale, multi-stakeholder collaborations to build technical expertise in patient-centered outcomes research, education, and dissemination. This type of input would get us much closer to determining value to a patient versus a single-outcome approach that assumes homogeneity of value across diverse populations. This would allow us to identify and address specific health disparities instead of obfuscating them, as happens in current value assessment methods.

Still other groups are designing an approach that realigns value assessment to patient risk reduction. A key component would be the addition of dynamic risk-scoring models to databases that would enable inclusion of risk and monitoring of changes in risk associated with key health status and health care indicators; social and environmental determinants of health; and access to quality health care, including guideline directed care.

One new approach developed by Cookson and colleagues⁴⁰ seeks to address health disparities directly and even allows for an empirical measure of the trade-off values between maximizing efficiencies and minimizing disparities. This has been used to evaluate how health policies in both high and low-income countries such as bowel cancer screening⁴¹ and rotavirus immunization⁴² may improve population health while at the same time exacerbating health disparities.

Conclusion

It is essential that we do not allow traditional value assessment that assess value to payers over patients to become embedded in how policymakers, payers and others assess the value of treatments. The issue described above is multi-faceted and complex. The solution will require an integrated, multi-faceted approach that will, in all likelihood, incorporate components of the innovative approaches described above and others that will surface as the conversation evolves. What is certain is that the development of an alternative to this system-wide inequity is essential.

34 "Developing Patient-Centered Models." Innovation and Value Initiative, 7 Jan. 2020, www.thevalueinitiative.org/developing-patient-centered-models/.

35 Baltussen R, Niessen L. Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost effectiveness and resource allocation*. 2006 Dec;4(1):14.

36 Zuniga MA, Carrillo-Zuniga G, Seol YH, Fos PJ. Multi-criteria assessment of county public health capability disparities. *Journal of health and human services administration*. 2009 Dec 1:238-58.

37 Marsh K, Dolan P, Kempster J, Lugon M. Prioritizing investments in public health: a multi-criteria decision analysis. *Journal of public health*. 2013 Sep 1;35(3):460-6.

38 Marsh K, Lanitis T, Neasham D, Orfanos P, Caro J. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. *Pharmacoeconomics*. 2014 Apr 1;32(4):345-65.

39 "Developing Patient-Centered Models." Innovation and Value Initiative, Innovation and Value Initiative , 7 Jan. 2020, www.thevalueinitiative.org/developing-patient-centered-models/.

40 Cookson R, Mirelman AJ, Griffin S, Asaria M, Dawkins B, Norheim OF, Verguet S, Culyer AJ. Using cost-effectiveness analysis to address health equity concerns. *Value in Health*. 2017 Feb 1;20(2):206-12.

41 Asaria, M., Griffin, S., Cookson, R., Whyte, S. and Tappenden, P., 2015. Distributional cost effectiveness analysis of health care programmes—a methodological case study of the UK bowel cancer screening programme. *Health economics*, 24(6), pp.742-754.

42 Dawkins, B.R., Mirelman, A.J., Asaria, M., Johansson, K.A. and Cookson, R.A., 2018. Distributional cost-effectiveness analysis in low-and middle-income countries: illustrative example of rotavirus vaccination in Ethiopia. *Health policy and planning*, 33(3), pp.456-463.

Endnotes

i A health utility is a numerical rating designed to represent health status and health-related quality of life.

ii Effect size references the amount of health improvement observed and measured from a given treatment