Manuscript Article File

Equity-informative methods of health services research

Structured Abstract

Purpose

We review quantitative methods for analysing the equity impacts of healthcare and public health interventions: who benefits most and who bears the largest burdens (opportunity costs)? Mainstream health services research focuses on effectiveness and efficiency but decision makers also need information about equity.

Approach

We review equity-informative methods of quantitative data analysis in three core areas of health services research: effectiveness analysis, cost-effectiveness analysis, and performance measurement. An appendix includes further readings and resources.

Findings

Researchers seeking to analyse health equity impacts now have a practical and flexible set of methods at their disposal which builds on the standard health services research toolkit. Some of the more advanced methods require specialised skills, but basic equity-informative methods can be used by any health services researcher with appropriate skills in the three core areas.

Value

We hope that this review article will raise awareness of equity-informative methods of health services research and facilitate their entry into the mainstream so that health policy makers are routinely presented with information about who gains and who loses from their decisions.

Keywords: (*only 6 allowed by the Emerald web submission system - suggested in yellow highlight; where possible have used terms suggested by the system e.g. Cost/Benefit Analysis, Inequality and terms that health researchers are likely to search for e.g. Subgroup Analysis rather than Conditional Average Treatment Effects.*

Conditional Average Treatment Effects, Cost/Benefit Analysis, Distributional Cost-Effectiveness Analysis, Equity, Inequality, Randomised Controlled Trials, Quality Improvement, Quality Indicators, Quasi Experimental Designs, Socioeconomic Factors, Small-Area analysis, Subgroup Analysis.

Article classification: General review

Plain Language Summary for "Kudos" platform:

This article reviews research methods for analysing the equity impacts of healthcare and public health interventions: who benefits most and who bears the largest burdens (opportunity costs)? More widespread use of these "equity-informative" methods could help decision makers make fairer decisions with better health outcomes.

1. Introduction

Healthcare and public health decision makers are increasingly concerned about inequalities between advantaged and disadvantaged groups in health, health service use, and financial hardship resulting from use of those services (for example, out-of-pocket costs) (Ottersen et al., 2014, Jamison et al., 2013, Ottersen et al., 2014, Marmot et al., 2008, World Health Organization, 2015). Systematic and substantial differences in health-related outcomes of this kind have been documented in relation to numerous equity-relevant variables, including social variables such as socioeconomic status, ethnicity and geographical location and disease categories such as disability and severe mental illness. Long-standing concerns about unfair health inequalities have been given further impetus in recent years by evidence of rising "deaths of despair" from suicide, drug overdose and alcoholic liver disease (Case and Deaton, 2020) and inequalities in coronavirus infection and mortality rates related to ethnicity (Galea and Abdalla, 2020).

Decision makers are increasingly motivated to address these concerns about health inequalities, but often poorly informed about the equity impacts of their decisions. This situation has arisen partly because different research disciplines often take narrow and contradictory approaches to evidence informed decision making. Mainstream health services research tends to prioritise effectiveness and efficiency over equity, focusing on average outcomes and providing little or no information about how those outcomes are distributed, i.e. who benefits most and who bears the largest burdens (opportunity costs). This runs the risk of adopting interventions that are effective on average but which widen inequalities. In contrast, health inequality researchers focus on equity rather than efficiency, emphasising variations in outcomes between groups, often assuming that reducing health inequalities will always generate net health gains for the whole population.

Both approaches miss important perspectives that the other provides. The consequences of this are unimportant if a policy decision is both efficient and equitable, i.e. if it improves overall

health and simultaneously reduces inequality. Unfortunately, improving equity may reduce efficiency or vice versa, and researchers and policy makers may be unaware of this problem. Health inequality researchers therefore need to join forces with health services researchers to find new ways of evaluating and monitoring health inequality solutions that explicitly address trade-offs between efficiency (improving total population health) and equity (reducing inequalities).

In that spirit, this article reviews some of the many "equity-informative" methods of health services research that now exist for analysing the health equity impacts of interventions and organisations. We focus on three core areas of health services research: effectiveness analysis, cost-effectiveness analysis, and performance measurement. We focus on introducing the basic concepts needed by policy advisers (e.g. health departments and reimbursement agencies) and research funders to understand and use equity-informative health services research, with further readings and resources for researchers in an appendix. Although some of the methods require specialised skills, simple and useful equity-informative methods can be applied by any health services researcher with skills in effectiveness analysis, cost-effectiveness analysis and/or performance measurement.

2. Equity-informative effectiveness analysis

The "gold standard" method of effectiveness analysis is the randomised controlled trial (RCT). However, quasi-experiments or natural experiments are also common – for example, regression discontinuity designs (RDD), interrupted time series (ITS) and difference-in-differences (DiD); the latter often combined with matching or weighting methods for selecting comparable control and treatment groups, such as propensity score matching and synthetic controls (Bedoya et al., 2017). Methods of systematic review and meta-analysis can then be used to combine effectiveness evidence from multiple effectiveness studies (O'Neill et al., 2014).

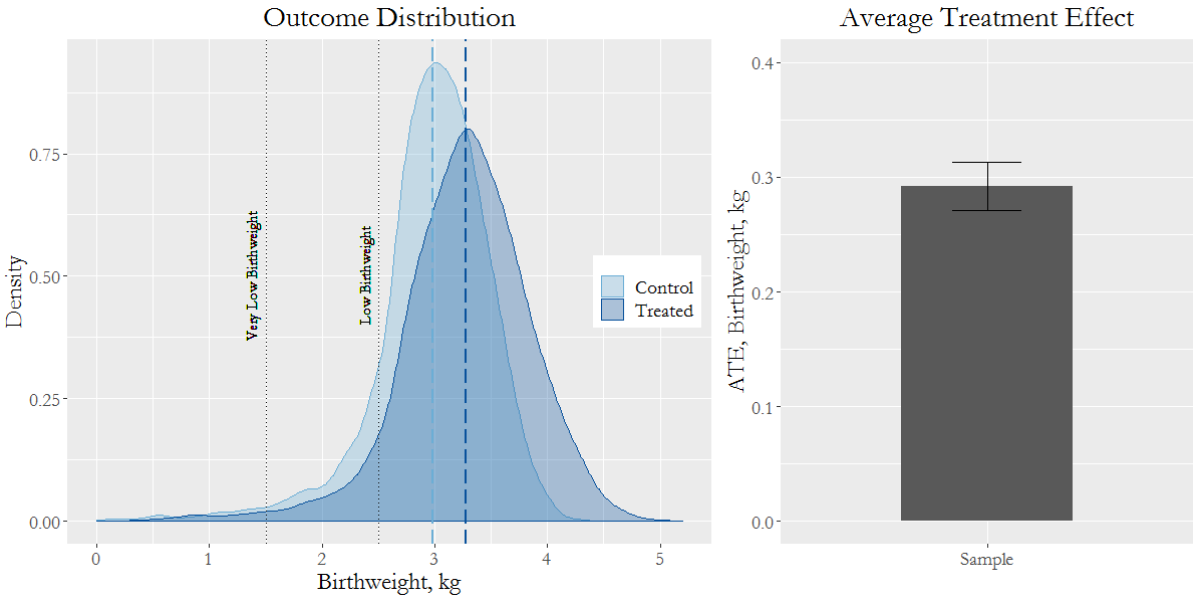
The purpose of all these approaches is to estimate "treatment effects" that measure the causal effect of an intervention on an outcome of interest. The intervention does not necessarily have to be a medical treatment (e.g. a drug, device or surgical procedure) – it might be a health prevention programme or indeed any decision option compared with an alternative. Typically, researchers estimate the Average Treatment Effect (ATE) which identifies the difference in the mean (average) outcome between the "treatment" group who receive the intervention and the "control" group who do not. In the special case of interventions targeted exclusively towards a disadvantaged population, the average treatment effect can provide some useful partial information about equity – it tells us whether, on average, the intervention benefits this disadvantaged population (Welch et al. 2017). More generally, however, average treatment effects fail to provide useful information about equity. This section reviews methods for providing such information – known as "Conditional Average Treatment Effects" (CATEs), "Quantile Treatment Effects" (QTEs) and "Inequality Treatment Effects" (ITEs). The estimation of CATEs is more commonly known as "subgroup analysis" in health services research, but we use the treatment effect terminology from the general literature on statistical methods of causal inference to help explain the relationship between these three different approaches.

2.1 Average Treatment Effect (ATE)

Figure 1 shows the effects on birthweight of a hypothetical antenatal dietary education and supplementation programme for a low-income country population. We use a control group distribution of birthweight that is typical for many low-income countries, with a mean birthweight of 2.98kg and a low birthweight (< 2.5kg) rate of 12.7% (Blencowe et al. 2019), and assume the programme delivers an average increase in birthweight of 0.29kg (Ota et al. 2015). Figure 1 shows a post-treatment follow-up, in which the intervention has been delivered to a 'treated' group of expectant mothers but not a 'control' group. The outcome of interest, y, is

birthweight. The distribution of birthweight within the control and treated groups is shown in the left panel, with mean birthweight shown by the dotted lines. The ATE is simply the difference between these averages: 0.29kg. This is shown in the right panel, with 95% confidence intervals.

Figure 1: Average Treatment Effect



Note: The left panel shows the distribution of birthweight for the control and treated groups. The dashed vertical lines show the mean birthweights for each group: 2.98kg for the control and 3.27kg for the treated. The dotted vertical lines show low (2.5kg) and very low (1.5kg) birthweights. The area under the curve to the left of these lines shows the percentage of low and very low birthweights: 12.7% and 1.8%, for the control, and 7.4% and 1.2% for the treated. The right panel shows the average treatment effect: 0.29kg, by the height of the bar, and 95% confidence intervals by the error bars.

However, the ATE ignores important effects on the distribution of birthweight. Below we illustrate methods to investigate this further.

2.2 Conditional Average Treatment Effects (aka Subgroup Analysis)

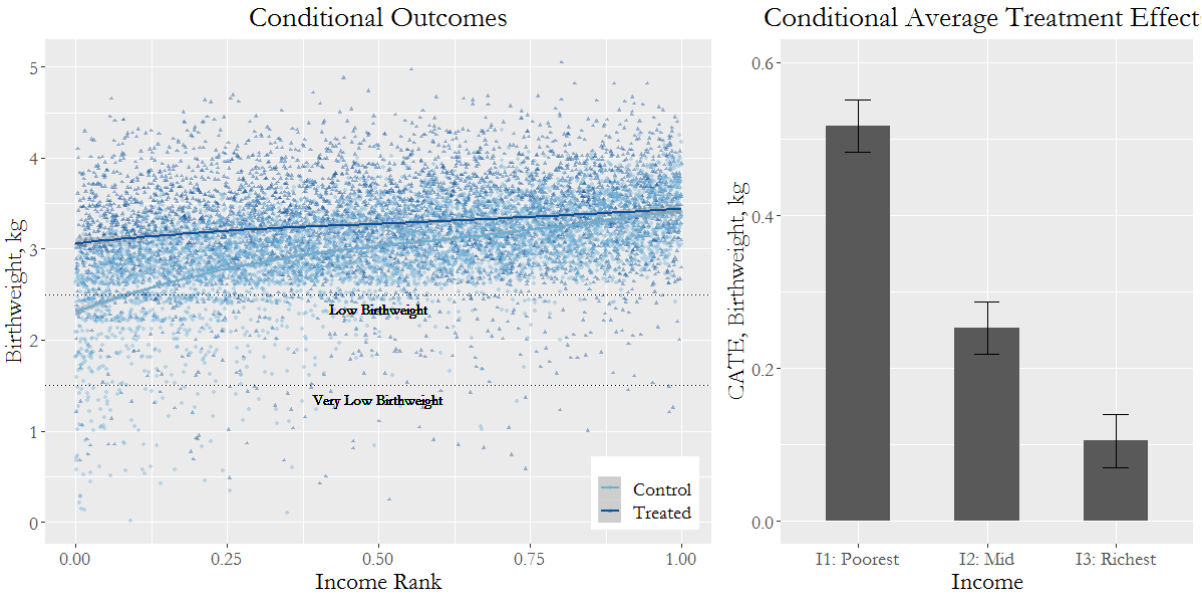
Decision makers are often concerned about inequalities in health outcomes associated with equity-relevant social variables (e.g. socioeconomic status, ethnicity, geographical location), disease categories (e.g. severity of illness, disability) and risk factors (e.g. blood pressure, obesity, smoking). We can estimate how intervention effects depend upon observed equity-relevant

variables using subgroup analysis or conditional average treatment effects (CATEs). The term 'conditional average treatment effects' is slightly more general than 'subgroup analysis', since the 'conditioning' variables of interest can include continuous variables like income as well as subgroups like income quantile group.

The left panel of Figure 2 shows how birthweight is related to household income rank, for both control and treated groups. Each dot represents one baby, with a particular birthweight and household income rank. For each group the mean birthweight, conditional on income rank, is shown by the blue lines. The relationship shown in the control group is typical: there is a steep 'social gradient' in health, whereby on average, babies with a higher income rank tend to have a larger birthweight. In the treated group, however, this social gradient is reduced, suggesting that babies from lower income households benefit more from the intervention than those from higher income households. The intervention is therefore pro-poor: it has greater benefits for poorer individuals.

To quantify these differential effects, researchers can estimate Conditional Average Treatment Effects (CATEs). CATEs show the expected difference in the outcome between control and treated group, conditional on observable characteristics. The right panel of Figure 2 shows CATEs estimated conditionally on belonging to one of three income subgroups: the poorest, middle and richest thirds of the population (i.e. 0.00-0.33 income rank 'poorest', 0.34 to 0.66 'middle', and 0.67 to 1.00 'richest'). The CATEs for each income rank are the vertical distances between the averaged lines in the left panel of Figure 2. The effects are largest on the poorest income group, with an expected increase of 0.52kg in birthweight. The richest group, on the other hand only sees an expected increase of 0.10kg. This effect is significantly lower than that of the poorest. The intervention has, therefore, significantly reduced health inequalities related to income. However, we do not know what happened to other kinds of inequality, for example inequality related to ethnicity, geographical location and gender.

Figure 2: Conditional Average Treatment Effects



Note: The left panel shows a scatter plot of household income rank against birthweight, for the control (circles) and treated (triangles) groups. The lines show the average birthweights conditional on income rank, for each group. The vertical difference between these lines are the conditional average treatment effects (CATEs). The right panel summarises the CATEs for the poorest, middle and richest third of the population, with 0.52kg, 0.25kg and 0.10kg, respectively. The 95% confidence intervals are shown by the error bars.

This example shows a simple subgroup analysis which splits a continuous variable (income) into three subgroups and uses dummy interaction terms to estimate the CATEs. But there are many ways of splitting a population into subgroups and groups are often defined arbitrarily, in terms of both number and boundaries. Different splits can lead to differences in the estimated CATEs and can be potentially misleading if they mask or exaggerate important features of the underlying relationship. More general methods include parametric interaction models, which interact the treatment effect with the conditioning variable using a specific functional form, and semi-parametric methods, which can estimate complex non-linear CATEs using locally-weighted regressions (Robson et al., 2019).

2.3 Quantile Treatment Effects

A different way of looking at equity impacts, which departs from the conventional focus on average health differences between socioeconomically disadvantaged groups, is to estimate

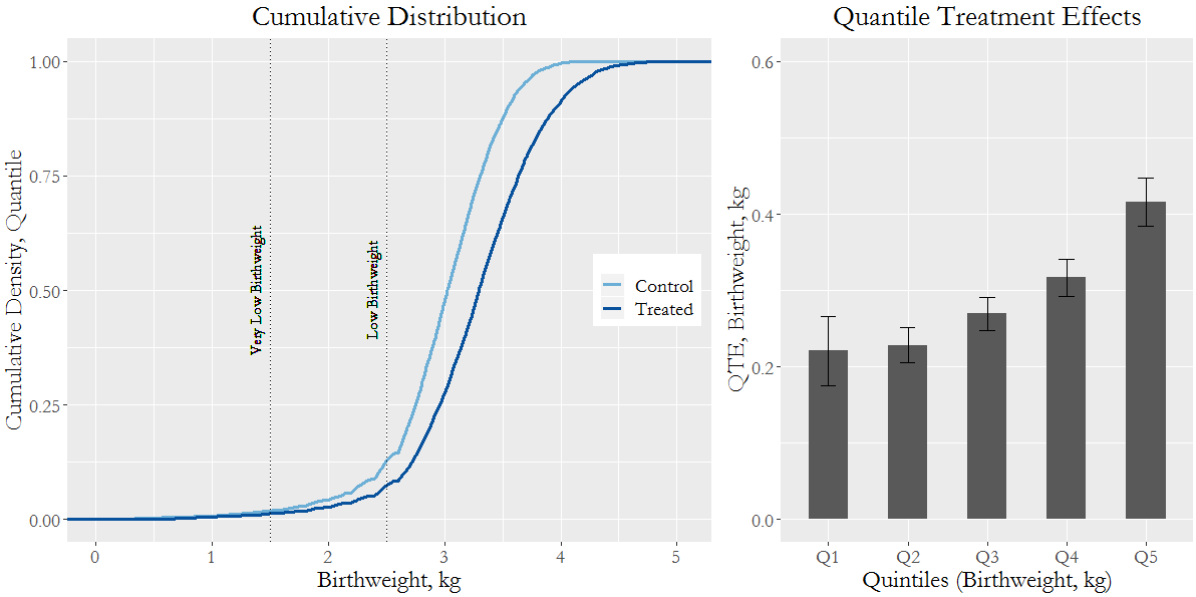
Quantile Treatment Effects (QTEs). Rather than focus on differences in treatment effects (i.e. impact on the health outcome) for different values of a predictor such as income, QTEs show the effect of the intervention at different quantiles of the health outcome.

In simple terms, CATEs tell us how far the intervention is 'pro-poor', whereas QTEs tell us how far the intervention is 'pro-unhealthy' – i.e. does it benefit recipients who are relatively unhealthy more than those who are relatively healthy? The two questions do not necessarily have the same answer. Although poorer groups may be healthier than richer ones on average, many poor individuals are healthy and many rich individuals are unhealthy (see Figure 2). Although this seems counter-intuitive, and may not happen often, it is possible for 'pro-poor' interventions to be 'pro-healthy' and thereby increase overall inequality in health. In relation to our hypothetical dietary programme, for example, women with supportive families might tend to have better birth outcomes in the control group and might also be better able to respond to antenatal dietary advice in the treatment group. The programme thus might have a stronger effect on birthweight not only in poorer families (an observed risk factor) but also in more supportive families (an unobserved risk factor).

Figure 3 illustrates the resulting 'pro-healthy' distribution of effects.[1] The right panel shows QTEs (i.e. gains in birthweight due to the intervention) for five health quantile points, where Quantile 3 (0.50) represents the effect at the middle of the health distribution (the median), Quantile 1 (0.10) the effect towards the bottom of the health distribution (the 10th percentile point), and Quantile 5 (0.90) the effect towards the top (the 90th percentile point). The left panel plots the cumulative distributions of the control and treated group, which gives the birthweight at a given quantile. The QTE is the horizontal difference between these curves. For Q3 (0.50) (the median) this difference is 0.27kg.

[1] Strictly speaking this is a 'pro-heavy' distribution, since over-weight babies are not necessarily healthier.

Figure 3: Quantile Treatment Effects



Note: The left panel shows the cumulative distribution curves of the control and treated groups. This shows the proportion of the population (quantile, q, on the y-axis) who have a birthweight (x-axis) below a certain level. The median birthweight (at q=0.5) is 3.02kg and 3.29kg for the control and treated group, respectively. At a given quantile, the horizontal difference between these two curves gives the quantile treatment effect (QTE). The right panel shows the QTEs at five quintiles, where the QTEs equal 0.22kg, 0.23kg, 0.27kg, 0.32kg and 0.42kg, for quantiles: 0.1, 0.3, 0.5, 0.7 and 0.9, respectively.

In this case, the effects are largest at the quantile with the highest birthweight, QTE(0.9) = 0.42kg, and smallest at the quantile with lowest birthweight, QTE(0.1) = 0.22kg. The intervention generates larger increases in birthweight at the higher end of the spectrum, thereby increasing 'pure' individual-level inequality in birthweight. Note that we are not comparing gains for individuals with different birthweights (babies are only born once); we are comparing the distributions of birthweights of babies born to mothers in the treated and control groups, finding greater differences between the groups at the higher end of the birthweight range. QTEs tell us about changes in the shape of the final distribution, not who precisely gains what.

2.4 Inequality Treatment Effects

A third approach is to estimate Inequality Treatment Effects (ITE), which compare the difference between summary indices of inequality within control and treatment groups. This approach is best used when the intervention is delivered across a general population, or a representative sample thereof, otherwise it will only measure inequality within a specific recipient population (e.g. a target age, disease, or risk factor group) and may under- or over-estimate the impact on the general population. For example, smoking cessation programmes usually achieve greater quit success among rich smokers than poor smokers, and so can appear to have a pro-rich inequality impact if one only looks within the subpopulation of smokers. However, there are far more poor smokers than rich smokers within the general population, and so smoking cessation programmes may have a pro-poor inequality impact within the general population.

Inequality indices use a single numerical value to represent an entire distribution and many indices are available (Kjellson et al. 2015). The ITE is the difference between these measures of inequality for the control and treated groups. In our above example, we use a Concentration Index to examine inequality in birthweight related to income, and a Gini Index to examine 'pure' individual-level inequality in birthweight. The Gini index is scaled from 0 to 1, with 0 representing full equality and 1 representing maximum inequality. For the Concentration Index, 0 represents full equality, positive values approaching 1 represent pro-rich inequality and negative values approaching -1 represent pro-poor inequality. The Concentration Index is 0.054 for the control and 0.017 for the treated, showing a 0.037 decrease in inequality related to income rank. The Gini Index is 0.090 for the control and 0.095 for the treated, showing a 0.005 increase in 'pure' inequality in birthweight. This provides confirmation that inequality in birthweight related to income has reduced while pure inequality in birthweight has increased.

2.5 Discussion

In practice, the biggest challenge when conducting equity-informative effectiveness analysis is data availability and sample size. Most RCTs are only designed and powered to estimate average treatment effects, and do not report information on equity-relevant covariates (Welch et al., 2017). There is a danger of spurious findings if one goes on a post-hoc 'fishing expedition' for equity effects (e.g. false negatives due to inadequate power, or false positives due to multiple comparisons) as there are numerous ways of splitting the data in search of apparently significant CATEs, QTEs or ITEs (Inglis et al., 2018, Burke et al. 2015). However, some RCTs do have large enough samples for equity analysis, as do many quasi-experimental studies based on population data, and sample size can be increased by individual-level pooling of data from many studies.

3. Equity-informative cost effectiveness analysis

Equity-informative effectiveness studies often fail to address wider equity issues of interest to decision makers, such as:

- who bears the health opportunity costs of diverting scarce resources,
- impacts on health inequality within the general population, beyond the study population
- long-term health inequality impacts, beyond the study follow-up period
- the size of health inequality impact compared with other programmes in other areas, and
- trade-offs between equity and efficiency objectives.

These issues can be addressed by equity-informative or 'distributional' cost-effectiveness analysis, which goes beyond standard cost-effectiveness analysis by using additional evidence and modelling to analyse equity impacts and trade-offs. Depending on the question in hand, the underpinning modelling can range from simple decision trees through to complex microsimulation (Skarda, I., Cookson, R., Asaria, M, 2020).

Distributional cost-effectiveness analysis (DCEA) is a broad umbrella term for studies that provide information about equity in the distribution of costs and effects as well as efficiency in terms of aggregate costs and effects (Cookson, R, Griffin, S, Norheim, OF. and Culyer, AJ, 2020). DCEA can involve simply exploring the implications of giving special priority or 'equity weight' to improving the health of programme recipients compared with non-recipients. It can also involve more detailed analysis of the distribution of health benefits and burdens within the general population by equity-relevant social variables (e.g. socioeconomic status, geographical location, indigenous status, ethnicity, gender, age), disease variables (e.g. disease classification, severity of illness, proximity to death, rarity of condition) or risk factors. It can also involve analysing distributional consequences for non-health outcomes, such as income or financial protection from

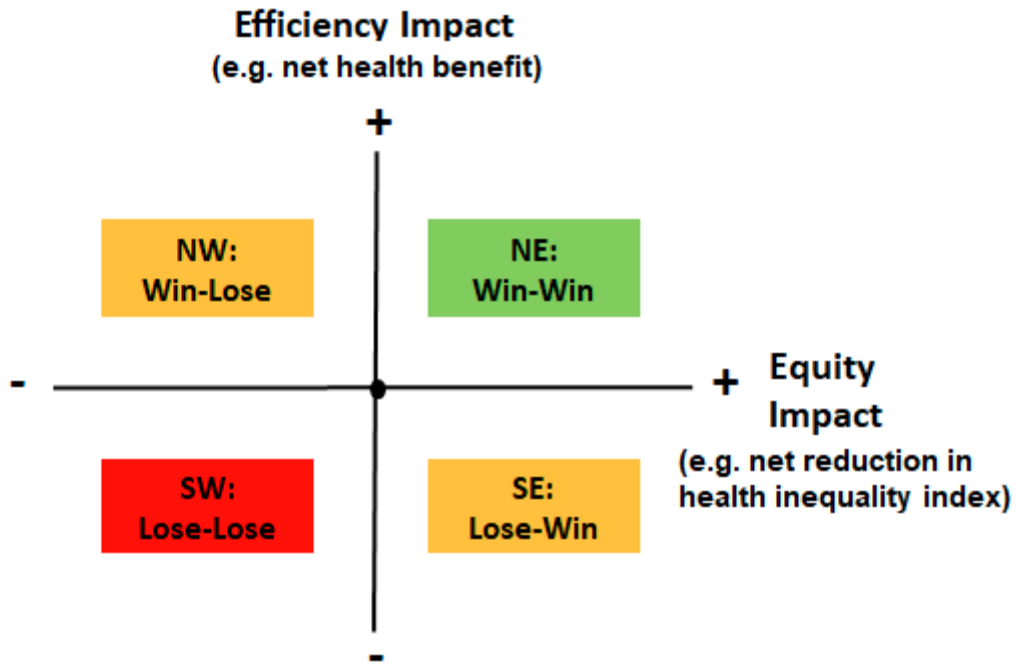
out-of-pocket health care costs, and evaluation of potential trade-offs between equity and efficiency objectives.

3.1 Equity-Efficiency Impact Plane

The equity-efficiency impact plane is a device to help analysts to think about trade-offs between efficiency and equity, and to consider whether more detailed equity analysis is necessary. It can also be used for visualising the findings of a DCEA study (Kypridemos et al., 2016) and to help decision makers keep both equity and efficiency objectives in sight.

A programme is cost-effective if its health benefit is greater than its health opportunity cost (Drummond et al., 2015). However, if it harms equity then a cost-effective programme might not be worth implementing. Conversely, if a cost-ineffective programme improves equity, then it might be worth implementing. The equity-efficiency impact plane in Figure 4 sets out the four logical possibilities in a systematic manner.

Figure 4: Equity-Efficiency Impact Plane



Both efficiency impact and equity impact can be measured in various ways, according to the objectives of the relevant decision maker.

A policy that falls in the 'win-win' quadrant improves both total health and health equity, and one that falls in the 'lose-lose' quadrant harms both. In low- and middle-income countries, vaccination and other infectious disease control programs often fall into the 'win-win' quadrant (Verguet et al., 2013, Verguet et al., 2015), as they typically deliver large health gains per unit cost and disproportionately benefit socially disadvantaged groups. By contrast, investments in high-cost treatments for late-stage chronic disease may fall into the 'lose-lose' quadrant of being neither cost-effective nor likely to reduce social inequality in health – in which case other ethical or political arguments are needed to justify funding (Asante et al., 2016).

Equity and efficiency impacts may also be opposed. In NW 'win-lose' quadrant, the option is cost-effective (i.e. is improves total health) but harms equity, and in the SE 'lose-win' quadrant,

the option is not cost-effective (i.e. is harms total health) but improves equity. When interventions fall in the 'win-lose' or 'lose-win' quadrant, it may be worth considering re-designing them to move closer to the 'win-win' quadrant. For example, if socially disadvantaged groups gain less than advantaged groups from a decision to fund a medical technology due to unequal access – placing the intervention in the 'win-lose' quadrant – additional investment to improve access for disadvantaged groups might improve the equity impact (i.e. shifting the intervention to the right on the equity impact axis) but it could also increase total programme costs (i.e. shifting the intervention downwards on the efficiency impact axis). Whether this re-design moves the intervention into the 'win-win' quadrant would need careful analysis.

3.2 Analysing Equity Impacts and Trade-Offs

In the 'win-lose' and 'lose-win' cases, equity trade-off analysis is required to address the questions: which policy is fairer, which policy is better for total health, and which policy is better overall? This analysis can be done informally by making intuitive judgements based on the pre- and post-policy distributions. It can also be done formally, for example by quantifying equity impacts using specific inequality indices or weighting benefits for certain groups. Training resources on how to apply these methods are provided in the appendix.

4. Equity-Informative Performance Measurement

Performance measurement is a core area of health services research and quality indicators are routinely used for quality improvement and accountability purposes (Smith et al., 2009). However, health inequality monitoring is often isolated from mainstream system-level, organisation-level and clinician-level healthcare performance monitoring and quality assurance processes. National reports on inequalities in health and healthcare are produced from time to time, but equity performance indicators are not routinely used to inform healthcare decision making or to hold healthcare planners, managers and clinicians accountable (Cookson et al.,

2018). As a result, health service planners, managers and clinicians remain ill-informed about inequalities in the quality of care within their own area of responsibility, despite becoming increasingly well-informed about quality of care for the average patient.

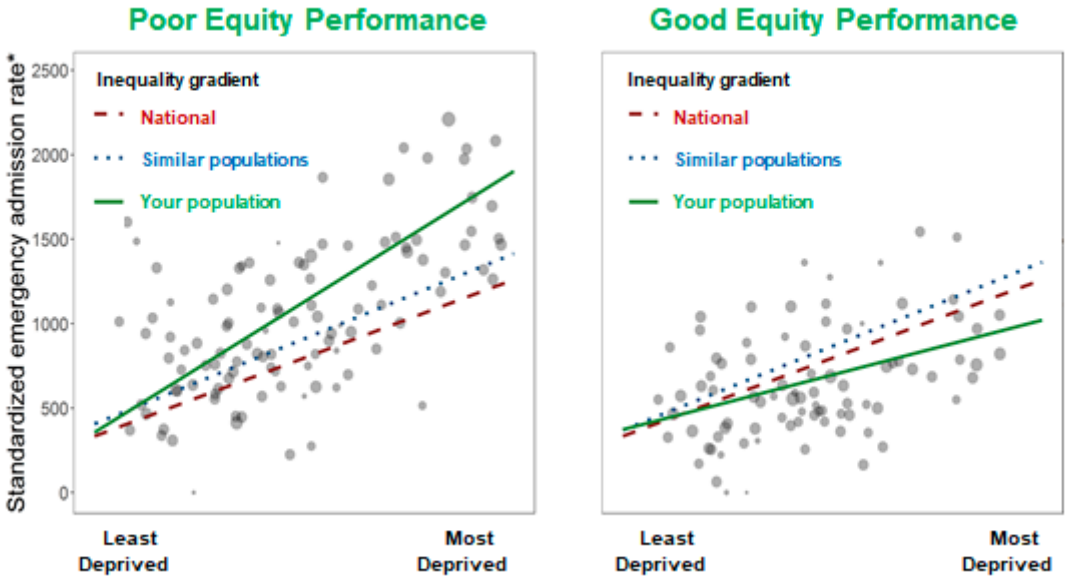
One problem has been the focus on 'sluggish' equity performance indicators like inequality in life expectancy, which are relatively unresponsive to short-term changes in health service delivery and are often several years out-of-date by the time they are reported. However, the major obstacle has been a lack of sufficiently granular analysis and reporting of equity performance. Indicators of inequality in health and healthcare have tended to be produced at the level of national or regional populations comprising millions of people. This allows inequality to be measured using a full range of standard quality indicators, including age-specific and disease-specific outcomes that are relatively rare within the general population, such as infant mortality or the recovery rate from psychological therapy. This can be useful for national and regional accountability purposes, especially in jurisdictions with a high degree of centralised control over national and regional health service planning. However, national and regional equity indicators do not speak directly to local health system planners responsible for healthcare purchasing, planning and delivery, or to the managers of hospitals, primary care practices and other health care provider organisations, or to clinicians themselves. They cannot be used to inform decision making by local planners, managers and clinicians or to hold them accountable.

A recent focus of research, therefore, has been to develop more granular equity performance indicators that are more relevant to local decision makers. At present, indicators exist that can be used for monitoring equity in healthcare within geographically-defined local health systems comprising populations of a few hundred thousand general population. This approach cannot be applied to age- and disease-specific quality indicators, since comparing different social subgroups then runs up against 'small numbers' problems of instability and lack of statistical power. However, it can be applied to several general indicators of healthcare structure, process

and outcome quality including – but not limited to – primary care supply, primary care process quality, hospital waiting times, hospital admissions, hospital mortality, and mortality considered amenable to health care (Cookson et al., 2016). The next phase of research will be to refine these indicators and develop new indicators for monitoring the equity performance of provider organizations and individual clinicians. Equity performance monitoring is still in its infancy, and equity indicators are not yet sufficiently well developed and understood to be used for 'performance management' and 'pay for performance' purposes involving high powered punishments and rewards. However, it is now possible to start using equity indicators for quality improvement and accountability purposes, to help monitor progress in improving equity, to spot potential emerging problems requiring further investigation, to learn lessons, and to provide a data analytical platform for evaluation studies.

In what follows, we present a general conceptual framework that can in principle be applied to equity performance comparisons at any population level. The basic idea is to measure social inequality in the quality of care within the population served by the relevant healthcare decision maker, benchmark this against similar populations served by similar decision makers, and monitor change over time. We illustrate the approach using an indicator that is widely used to assess the quality of care coordination between primary, community and acute settings: emergency inpatient hospital admissions for ambulatory care sensitive chronic conditions ('potentially avoidable emergency admissions', for short). Many other care quality indicators can be used, but this is a useful headline indicator for national, regional and local health system monitoring that can be measured using readily available administrative data on hospital activity and updated annually (or quarterly with an annual moving average). The NHS in England has used an equity indicator of this kind since 2016 (Cookson et al., 2016). The basic conceptual framework is illustrated in Figure 5.

Figure 5: Equity Performance Measurement: Conceptual Framework



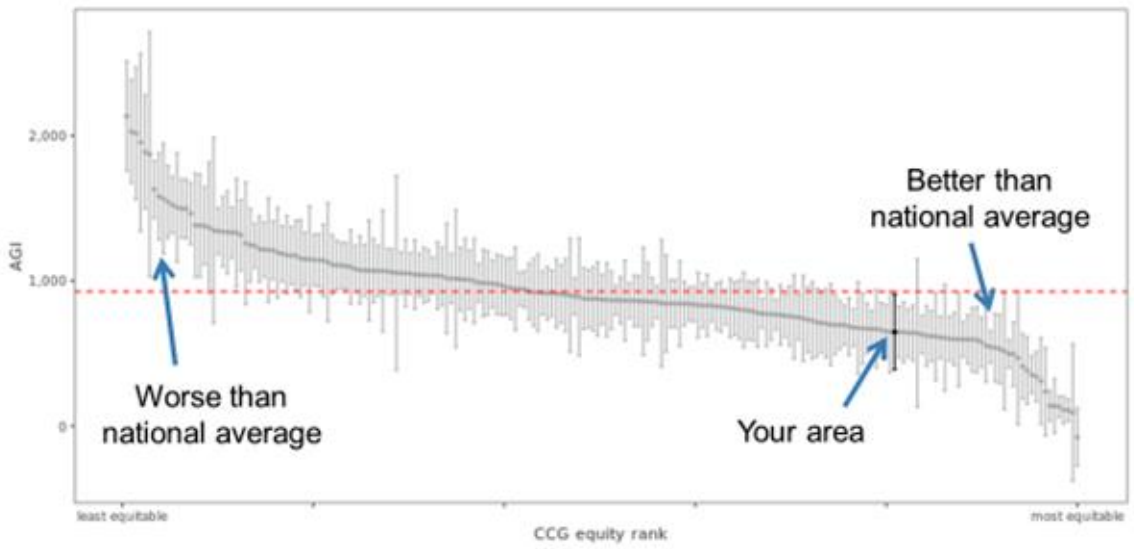
Note: Rate of emergency admissions for ambulatory care sensitive conditions (e.g. heart and lung disease, diabetes, dementia) in each small area neighbourhood within the local health system, indirectly age-sex adjusted. The assumption is these admissions are potentially avoidable through high quality care, and hence more is worse – though of course some emergency admissions are unavoidable and even desirable. Source: https://shiny.york.ac.uk/ccg_equity/

Equity performance is compared between the population served by the relevant healthcare decision making unit – labelled "your" population in Figure 6 – and similar benchmark populations. Benchmarking against similar populations is crucial, because some of the observed inequality in healthcare quality and outcomes may be due to underlying social variation in health risk factors outside the control of the healthcare decision making unit. In this example, the healthcare decision making units are 209 health care planning areas in England ("clinical commissioning groups") with a mean population of 272,000. Inequality is measured by analysing social variation in care quality within the relevant population – in this example, small-area variation between neighbourhoods with a mean population of 1,500. An index of inequality is estimated based on the association between care quality and a nationally comparable index of

social disadvantage – in this case, the English index of multiple deprivation, which is analogous to the Australian Socio-Economic Indexes for Areas. The use of a nationally comparable index of social disadvantage is crucial, because it means that estimated inequality gradients are then comparable between different sub-national populations with different mixtures of advantaged and disadvantaged people. Sub-national gradients can then be thought of as modelled estimates of the national gradient, if the nation had the same distribution of health outcomes as your population. We highlight this by using non-standard labels for sub-national inequality indices – for example, we use the term 'Absolute Gradient Index' (AGI) to denote the sub-national estimate of the national 'Slope Index of Inequality' (SII). Equity performance can then be assessed by comparing inequality in your population with two performance benchmarks: inequality in 'similar' populations and national inequality. The 'similar' populations can be selected using algorithms based on population characteristics, or by asking the healthcare decision maker to select their own comparison populations.

Equity performance can then be compared in cross section between different healthcare decision making units, for example using caterpillar plots of the kind illustrated in Figure 7.

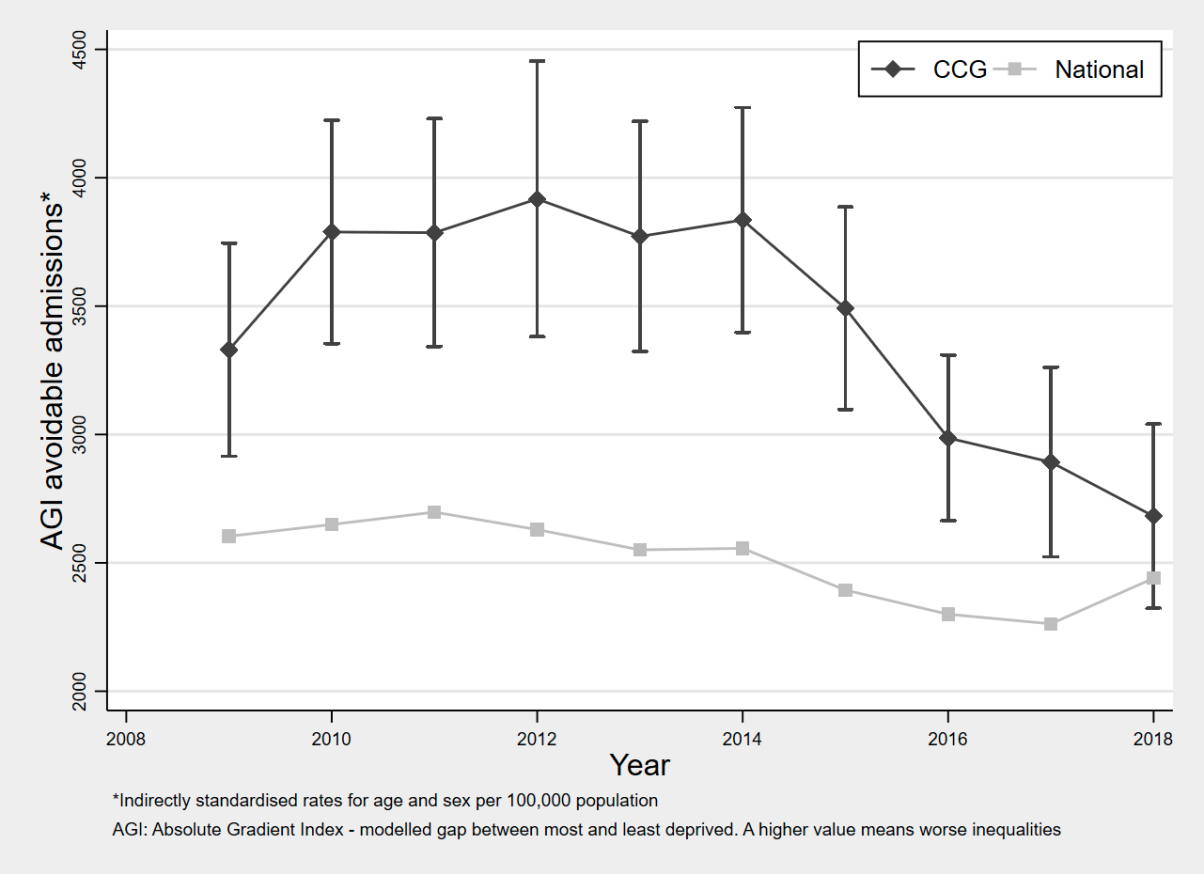
Figure 6: Equity Performance Comparison, With 95% Confidence Intervals



Note: Comparison between 209 English local health planning areas in 2015, based the Absolute Gradient Index (AGI) of deprivation-related inequality in avoidable emergency hospitalisation. As before, a higher rate of potentially avoidable emergency admissions is presumed to indicate worse quality ambulatory care, so long as one is comparing similar populations with similar levels of morbidity and risk factors. Source: https://shiny.york.ac.uk/ccg_equity/

Equity performance can also be compared over time for the same decision-making unit using time series plots of the kind illustrated in Figure 7.

Figure 7: Equity Performance Trends, With 95% Confidence Intervals



Note: This shows equity performance trends for a single local NHS planning area in England, from 2008 to 2018.

Source: https://www.york.ac.uk/che/research/equity/monitoring/packs/

Equity trend data of this kind can then be used to learn quality improvement lessons, by conducting case studies of decision-making units which have shown sustained improvement or worsening of equity over time. As with all performance measurement exercises, however, caution must be taken as there is a risk of spurious findings due to data artefacts and multiple comparisons. Administrative data are vulnerable to bias and instability due to coding errors and changes in data collection processes, and it is not possible reliably to estimate social gradients within populations with limited social variation – for example, populations in which almost everyone is either rich or poor. Furthermore, with a 95% confidence interval we would expect 1 in 20 decision making units to show abnormal equity performance. In the case of equity performance, the scope for multiple comparisons is increased further by the multiplicity of

different inequality indices that are available. For example, as well as the AGI index one could estimate its relative inequality counterpart, the 'Relative Gradient Index', or one could look at simple gap indices such as the difference between the top and bottom social group, or indeed any number of other weird and wonderful inequality indices that are available (Kjellsson et al., 2015). However, if a population shows a clear, sustained, and statistically significant improvement or worsening in an equity performance indicator there is at least a *prima facie* case for further investigation.

Where suitably granular and high-quality data are available, this approach could be extended to other indicators of social disadvantage – such as gender, ethnicity and rurality – and could be applied to data on individual-level variations in care quality as well as small-area variations. Current applications have focused on equity performance comparisons between the general populations served by geographically-defined national, regional or local health system planners. In principle, however, this approach could be used for equity performance comparisons between the specific populations served by health plans, provider organisations and individual clinicians. This will be essential in fragmented health systems without a 'single payer' responsible for general population health coverage, and is an important challenge even within single payer systems since hospitals and other provider organisations have considerable power and influence and equity improvement is not possible without their cooperation. Comparing equity performance between non-general populations makes it more difficult to select appropriate 'similar' populations for equity benchmark comparisons. and raises problems of case-mix adjustment that have not yet been thoroughly addressed and are key challenges for future research in this area.

5. Discussion

Increasing awareness of health inequalities has placed additional pressure on providers of health and related services to deliver care that is not only safe and effective, but also equitable. A wealth of research demonstrates that wide inequalities in access and outcomes remain between geographically and socially distinct groups, even in universal healthcare systems. Although they have limited control of the wider social determinants of health inequalities, providers are increasingly recognizing that they must do more to assess and address the equity impact of their own interventions.

In some respects, the movement towards equity-informative evidence mirrors the wider quality and effectiveness movement in medicine. For much of the twentieth century, the medical profession – lacking access to routine activity and outcomes data and comparative methods of analysis – remained in denial about variations in quality. Extreme examples of negligent care and confidential enquiries brought this stance into question, but it was not until data were routinely collected and analysed across whole systems that the scale of suboptimal care – and potential solutions to it – became apparent (Institute of Medicine and Committee on Quality of Health Care in America, 2001). The medical and related professions have since put systems in place for measuring and improving safety and effectiveness, with varying degrees of success.

The effectiveness movement initially focused on individual care, as this is the primary responsibility of care providers and reflects the traditional emphasis of their training. More recently, however, there has been a reorientation from focusing on the causes of individual cases to addressing the causes of population incidence (Rose, 2001). Definitions of quality have been expanded to include population-level considerations, principally efficiency, requiring more data and new methods. The development of measures such as QALYs (quality-adjusted life years) has allowed for meaningful comparisons across interventions, but from the outset these methods

were criticised for ignoring variation across patient and population groups (Sassi et al., 2001). The spread of these new methods coupled with a failure to look beyond averages has created the risk that differences in the context of targeted populations will be overlooked and that well-intentioned public health interventions will worsen health inequalities even as they improve efficiency. For example, the UN's Millennium Development Goals had the explicit goal of reducing inequalities within and between countries but relied on aggregate measurements of overall progress, ignoring inequitable distributions (Friedman et al., 2019). This has led to pleas for greater disaggregation of data. However, using old tools on these new data is likely to perpetuate the problem; we also need novel methods that are more equity informative.

In this paper, we have described equity-informative methods that build on standard health services research tools to provide practical, flexible and powerful ways of analysing health equity impacts and trade-offs. These methods can be used to inform decisions by providers and policy makers about the funding, design and delivery of interventions expected to have different consequences for different people, including decisions about health technology purchasing, benefit package design, organisation and delivery of care, and investment in prevention. The methods can be also applied by health services researchers with appropriate skills in the core areas of effectiveness analysis, cost-effectiveness analysis or performance measurement. However, such activity requires support. Policy makers and research funders need to play their part in reshaping the health services research infrastructure in pursuit of fairer decisions and better health by funding and adopting more equity-informative methods. Using equity-informative methods routinely alongside conventional health services research methods will then ensure that decision making is routinely informed by evidence not only about effectiveness and efficiency but also about equity impacts and trade-offs.

Appendix

Resources for equity-informative effectiveness analysis

Further readings

Bedoya, Bittarello, L, Davis, J, Mittag. N. (2017), *Distributional Impact Analysis: Toolkit and Illustrations of Impacts Beyond the Average Treatment Effect,* The World Bank, available at:https://doi.org/10.1596/1813-9450-8139.

Firpo, S. and Pinto, C. (2016), "Identification and estimation of distributional impacts of interventions using changes in inequality measures", *Journal of Applied Econometrics*, Wiley Online Library, Vol. 31 No. 3, pp. 457–486.

Robson, M., Doran, T., Cookson, R. and Others. (2019), Estimating and Decomposing Conditional Average Treatment Effects: The Smoking Ban in England, HEDG Working Paper, Department of Economics, University of York, available at: https://www.york.ac.uk/media/economics/documents/hedg/workingpapers/1920.pdf.

Sun X, Ioannidis JP, Agoritsas T, Alba AC, Guyatt G. How to use a subgroup analysis: users' guide to the medical literature. Jama. 2014;311(4):405-11.

Welch, V.A., Norheim, O.F., Jull, J., Cookson, R., Sommerfelt, H., Tugwell, P. and CONSORT-Equity and Boston Equity Symposium. (2017), "CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials", *British Medical Journal*, Vol. 359, p. j5085.

Other resources

To aid applied researchers useful Stata commands for estimating CATEs, QTEs and ITEs are shown below. We denote the outcome, *y*, the treatment group indicator variable, *treat*, and the conditioning variable (e.g. socioeconomic status), *x1*. We provide the simplest forms of the commands used for analysing RCTs; appropriate modifications may be needed for analysing quasi-experimental designs, such as regression discontinuity and difference-in-difference, and using sample weights or propensity score matching.

reg y treat – estimate ATE

reg y treat x1 treat#x1 - estimate simple CATEs, using interaction term approach

lwcate reg treat, xtilde(x1) – estimate complex non-linear CATEs using locally weighted regressions, for continuous, discrete and multidimensional conditioning variables

qreg y treat – estimates QTEs, at specified quantiles

grqreg treat – plots QTEs across all quantiles

conindex y , compare(treat) – Gini Coefficients estimated, and ITEs shown

conindex y , compare(treat) rank(income) – Concentration Index, and bivariate ITEs

Resources for equity-informative cost-effectiveness analysis

Further readings

Cookson, R, Griffin, S, Norheim, OF. and Culyer, AJ (Eds.). (2020), *Distributional Cost-Effectiveness Analysis: quantifying Health Equity Impacts and Trade-Offs*, Oxford University Press.

Cookson, R., Mirelman, A.J., Griffin, S., Asaria, M., Dawkins, B., Norheim, O.F., Verguet, S., et al. (2017), "Using Cost-Effectiveness Analysis to Address Health Equity Concerns", *Value in Health:* Vol. 20 No. 2, pp. 206–212.

Verguet, S., Laxminarayan, R. and Jamison, D.T. (2015), "Universal public finance of tuberculosis treatment in India: an extended cost-effectiveness analysis", *Health Economics*, Vol. 24 No. 3, pp. 318–332.

Verguet, S., Kim, J.J. and Jamison, D.T. (2016), "Extended Cost-Effectiveness Analysis for Health Policy Assessment: A Tutorial", *PharmacoEconomics,* Vol. 34 No. 9, pp. 913–923.

Other resources

- International Health Economics Association Special Interest Group on Equity-Informative Economic Evaluation https://www.healtheconomics.org/page/EEE

- Distributional cost-effectiveness analysis resources https://www.york.ac.uk/che/research/equity/distributional-cost-effectiveness-analysis/

- Freely downloadable spreadsheet training exercises in distributional cost-effectiveness analysis, accompanying the Oxford University Press handbook: https://www.york.ac.uk/che/research/equity/handbook/

- Online distributional cost-effectiveness analysis tool: https://shiny.york.ac.uk/dcea/.

Resources for equity-informative performance measurement

Further readings

Cookson, R., Asaria, M., Ali, S., Shaw, R., Doran, T. and Goldblatt, P. (2018), "Health equity monitoring for healthcare quality assurance", *Social Science & Medicine*, Vol. 198, pp. 148–156.

Cookson, R., Asaria, M., Ali, S., Ferguson, B., Fleetcroft, R., Goddard, M., Goldblatt, P., et al. (2016), Health Equity Indicators for the English NHS: A Longitudinal Whole-Population Study at the Small-Area Level, NIHR Journals Library, Southampton (UK), available at: https://doi.org/10.3310/hsdr04260

Other resources

Local equity performance measurement

- Local health equity indicators for the NHS in England (University of York) https://www.york.ac.uk/che/research/equity/monitoring/
- Local equity indicator data packs for England (University of York) https://www.york.ac.uk/che/research/equity/monitoring/packs/
- Reducing health inequalities resources (NHS England) https://www.england.nhs.uk/about/equality/equality-hub/resources/
- Health Inequality NHS Right Care Packs (NHS England) https://www.england.nhs.uk/rightcare/products/ccg-data-packs/equality-and-health-inequality-nhs-rightcare-packs/
- Reducing health inequalities (Public Health Scotland, including local health and social care inequality indicators within the ScotPHO profiling tool) http://www.healthscotland.scot/reducing-health-inequalities/use-the-right-indicators

National and regional equity performance measurement

- Handbook of health inequality monitoring with a special focus on low- and middle-income countries (WHO) https://www.who.int/gho/health_equity/handbook/en/
- National Healthcare Quality and Disparities Reports (US Agency for Healthcare Research and Quality) https://www.ahrq.gov/research/findings/nhqrdr/index.html
- Health inequalities (Canadian Institute for Health Information) https://www.cihi.ca/en/health-inequalities

References

Asante, A., Price, J., Hayen, A., Jan, S. and Wiseman, V. (2016), "Equity in Health Care Financing in Low- and Middle-Income Countries: A Systematic Review of Evidence from Studies Using Benefit and Financing Incidence Analyses", *PloS One*, Vol. 11 No. 4, p. e0152866.

Bedoya, G., Bittarello, L., Davis, J. and Mittag, N. (2017), *Distributional Impact Analysis: Toolkit and Illustrations of Impacts Beyond the Average Treatment Effect*, The World Bank, available at:https://doi.org/10.1596/1813-9450-8139.

Blencowe H, Krasevec J, de Onis M, et al. National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *The Lancet Global Health*. 2019; 7: e849-e60.

Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. *British Medical Journal.* 2015; 351:h5651.

Case, A. and Deaton, A. (2020), *Deaths of Despair and the Future of Capitalism*, Princeton University Press, available at: https://play.google.com/store/books/details?id=anK0DwAAQBAJ.

Cookson, R., Asaria, M., Ali, S., Ferguson, B., Fleetcroft, R., Goddard, M., Goldblatt, P., et al. (2016), *Health Equity Indicators for the English NHS: A Longitudinal Whole-Population Study at the Small-Area Level*, NIHR Journals Library, Southampton (UK), available at:https://doi.org/10.3310/hsdr04260.

Cookson, R., Asaria, M., Ali, S., Shaw, R., Doran, T. and Goldblatt, P. (2018), "Health equity monitoring for healthcare quality assurance", *Social Science & Medicine*, Vol. 198, pp. 148–156.

Cookson, R, Griffin, S, Norheim, OF. and Culyer, AJ (Ed.). (2020), *Distributional Cost-Effectiveness Analysis:quantifying Health Equity Impacts and Trade-Offs*, Oxford University Press.

Drummond, M.F., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W. (2015), *Methods for the Economic Evaluation of Health Care Programmes*, Oxford University Press, available at: https://play.google.com/store/books/details?id=yzZSCwAAQBAJ.

Friedman, E.A., Gostin, L.O., Kavanagh, M.M., Periago, M.R., Marmot, M., Coates, A., Binagwaho, A., et al. (2019), "Putting health equity at heart of universal coverage—the need for national programmes of action", *British Medical Journal*, Vol. 367, available at:https://doi.org/10.1136/bmj.l5901.

Galea, S. and Abdalla, S.M. (2020), "COVID-19 Pandemic, Unemployment, and Civil Unrest: Underlying Deep Racial and Socioeconomic Divides", *JAMA: The Journal of the American Medical Association*, available at:https://doi.org/10.1001/jama.2020.11132.

Griffin, S., Love-Koh, J., Pennington, B. and Owen, L. (2019), "Evaluation of Intervention Impact on Health Inequality for Resource Allocation", *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, Vol. 39 No. 3, pp. 171–182.

Inglis, G., Archibald, D., Doi, L., Laird, Y., Malden, S., Marryat, L., McAteer, J., et al. (2018), "Credibility of subgroup analyses by socioeconomic status in public health intervention evaluations: An underappreciated problem?", *SSM - Population Health*, Vol. 6, pp. 245–251.

Institute of Medicine and Committee on Quality of Health Care in America. (2001), *Crossing the Quality Chasm: A New Health System for the 21st Century*, National Academies Press, available at: https://play.google.com/store/books/details?id=OJ1FcYY3htAC.

Jamison, D.T., Summers, L.H., Alleyne, G., Arrow, K.J., Berkley, S., Binagwaho, A., Bustreo, F., et al. (2013), "Global health 2035: a world converging within a generation", *The Lancet*, Vol. 382 No. 9908, pp. 1898–1955.

Kjellsson, G., Gerdtham, U.-G. and Petrie, D. (2015), "Lies, Damned Lies, and Health Inequality Measurements: Understanding the Value Judgments", *Epidemiology* , Vol. 26 No. 5, pp. 673–680.

Kypridemos, C., Allen, K., Hickey, G.L., Guzman-Castillo, M., Bandosz, P., Buchan, I., Capewell, S., et al. (2016), "Cardiovascular screening to reduce the burden from cardiovascular disease: microsimulation study to quantify policy options", *BMJ* , Vol. 353, p. i2793.

Marmot, M., Friel, S., Bell, R., Houweling, T.A.J., Taylor, S. and Commission on Social Determinants of Health. (2008), "Closing the gap in a generation: health equity through action on the social determinants of health", *The Lancet*, Vol. 372 No. 9650, pp. 1661–1669.

Ota E, Hori H, Mori R, et al. Antenatal dietary education and supplementation to increase energy and protein intake. *The Cochrane database of systematic reviews*. 2015: Cd000032.

O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., Evans, T., et al. (2014), "Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health", *Journal of Clinical Epidemiology*, Vol. 67 No. 1, pp. 56–64.

Ottersen, T., Norheim, O.F. and World Health Organization Consultative Group on Equity and Universal Health Coverage. (2014), "Making fair choices on the path to universal health coverage", *Bulletin of the World Health Organization*, Vol. 92 No. 6, p. 389.

Robson, M., Doran, T., Cookson, R. and Others. (2019), *Estimating and Decomposing Conditional Average Treatment Effects: The Smoking Ban in England*, HEDG, c/o Department of Economics, University of York, available at: https://www.york.ac.uk/media/economics/documents/hedg/workingpapers/1920.pdf.

Sassi, F., Archard, L. and Le Grand, J. (2001), "Equity and the economic evaluation of healthcare", *Health Technology Assessment* , Vol. 5 No. 3, pp. 1–138.

Shah, K.K., Cookson, R., Culyer, A.J. and Littlejohns, P. (2013), "NICE's social value judgements about equity in health and health care", *Health Economics, Policy, and Law*, Vol. 8 No. 2, pp. 145–165.

Skarda, I., Cookson, R., Asaria, M. (2020), "Lifecourse Modelling for Childhood Policy Analysis", available at: https://equipol.org/research/projects/lifesim/ (accessed 14 July 2020).

Smith, P.C., Mossialos, E., Leatherman, S. and Papanicolas, I. (2009), *Performance Measurement for Health System Improvement: Experiences, Challenges and Prospects*, Cambridge University Press,

available at: https://play.google.com/store/books/details?id=c4CsWZYVOTwC.

Verguet, S., Laxminarayan, R. and Jamison, D.T. (2015), "Universal public finance of tuberculosis treatment in India: an extended cost-effectiveness analysis", *Health Economics*, Vol. 24 No. 3, pp. 318–332.

Verguet, S., Murphy, S., Anderson, B., Johansson, K.A., Glass, R. and Rheingans, R. (2013), "Public finance of rotavirus vaccination in India and Ethiopia: an extended cost-effectiveness analysis", *Vaccine*, Vol. 31 No. 42, pp. 4902–4910.

Welch, V.A., Norheim, O.F., Jull, J., Cookson, R., Sommerfelt, H., Tugwell, P. and CONSORT-Equity and Boston Equity Symposium. (2017), "CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials", *British Medical Journal* , Vol. 359, p. j5085.

World Health Organization. (2015), *Tracking Universal Health Coverage: First Global Monitoring Report*, World Health Organization, available at: https://play.google.com/store/books/details?id=d140DgAAQBAJ.