Towards theoretically robust evidence on health equity: A systematic approach to contextualizing equity-relevant randomized controlled trials

INTRODUCTION

Reducing inequalities in health and the determinants of health is a widely acknowledged health policy goal [1, 2], and methods for measuring inequalities and inequities in health are well developed [3-5]. Yet, the evidence base is weak for how to achieve these goals. There is a lack of high-quality randomized controlled trials reporting impact on the distribution of health and non-health benefits, and lack of methodological rigour in how to design, power, measure, analyse, and interpret distributional impact in randomized controlled trials (RCTs) [6, 7].

One reason is that this is a new and growing field of research and standards are not yet well established. Another reason is that equality and equity are inherently complex notions [8], and lack of clarity about which underlying health equity framework, if any, is used as basis for choosing inequality measures may hamper agreement and standardisation. Recent work has begun to address these issues: Welch et al’s CONSORT-Equity extension proposes an approach for standardized and systematic reporting of health equity in randomized trials [9], and Jull et al have proposed a consensus-based approach to conceptualizing health-equity relevant RCTs [10]. However, there is more scope for discussing how equity relevance in RCTs can be linked more explicitly to broader theories of health equity. Different conceptions of health equity will have different implications for the kinds of studies that should be undertaken and for the analysis and interpretation of the results.

Our overarching aim in this paper is to contribute to this emerging effort towards greater transparency and coherence in the theoretical and conceptual basis for RCTs on effective interventions to reduce health inequity. We endeavour to achieve this aim by pursuing two more specific objectives. First, we provide an overview of three major accounts of health equity in the literature, and clarify their implications for the measurement of health inequality in RCTs. Second, we seek to clarify the relationship between theory and translational challenges that researchers would need to attend to, in order to ensure that equity relevant RCTs are coherently grounded in theory.
We describe how to translate theory into practice in a consistent manner across the different stages of planning, conducting, and reporting the results of an equity-relevant RCT. We also outline some further concerns for researchers to address.

**WHAT IS HEALTH EQUITY?**

One key issue in debates about health equity in the philosophical and public health literature, is the question *which* inequalities in health are inequitable, or unjust.¹ There is a general consensus that not all inequalities are unjust, so we need an account of which inequalities are. Different theorists give different answers to that question. In the following, we propose an overview of three types of accounts of health equity. Our overview is not exhaustive, and some views may combine elements from the different categories. Furthermore, we do not here wish to argue in favour of one account over any of the others.²

Any account of health equity will, whether implicitly or explicitly, incorporate certain assumptions about the more fundamental question of how justice in health is related to distributive justice more broadly. Most plausibly, health is but one of several ‘currencies’ of justice within a pluralist framework, where justice concerns the distribution of more than one good (for example, health, income, and education). One underlying question, then, is the extent to which fairness in the distribution of health is contingent on the fairness of our social arrangements and the distribution of other justice-relevant goods. Following Peter [14], we can characterise accounts of health equity as ‘direct’ or ‘indirect’, depending on what position they take on this issue.

On the direct approach, the distribution of health can in itself be just or unjust, independently of how other important social goods, such as income, are distributed. As Peter explains, the distribution of health is considered in isolation from other social spheres or goods, and health equity ‘is defined with respect to particular distributions of health outcomes’ [14]. One might take this view if one considers health to be ‘special’, or of greater moral importance than other constituents of overall well-being, such that we require a separate standard or principle of fairness for health. Anand [15], for example, argues that inequalities in health are generally

---

¹ We acknowledge that in drawing exclusively on the philosophical and public health literature on health inequality and inequity, we have left out other relevant and valuable perspectives, e.g. in the sociological literature on social class and inequality (Bourdieu [11, 12], Giddens [13]).

² In contrast, Welch et al [9] adopt a specific conception of health equity as the basis for their proposed CONSORT-Equity extension.
morally worse than inequalities in other goods, and that therefore we should be less accepting of inequality in the distribution of health. Needless to say, one can give different answers with respect to what level of inequality is acceptable. In addition, one might think that considering health inequalities through the lens of established theories of distributive justice might exclude potentially unfair health inequalities. Christopher Murray [16], who advocates taking a comprehensive approach to health inequality, represents this view:

“Braveman et al believe that health inequalities correlated with factors other than income, social class, and race are not morally important. [...] For a child with an increased risk of death because she lives in a community with a poor immunisation programme and a high prevalence of HIV, it is no solace to know that her risk of death is uncorrelated with income, social class, or race. To most of us, inequality is the state of being unequal. Health inequalities exist when individuals' risks of death and poor health are unequal.”

On an indirect account of health equity, justice in health is interrelated with distributive justice more broadly. Whether or not a distribution of health is fair is contingent, in one way or another, on the distribution of other relevant goods. Here, we introduce a further distinction between two different (but not mutually exclusive) versions of this approach. The first version, most explicitly defended by Daniel Hausman [17, 18], conceives the correct ‘space’ of distributive justice as the distribution of overall well-being or advantage.iii We can call this the ‘overall disadvantage’ approach. The idea here is that the different components that make up well-being are to some extent substitutable, so that having more of one component can, partly or fully, make up for having less than another.iv This view of justice implies that, bar extremes, there is no particular distributive requirement for each of the relevant goods, or components of well-being, on its own. The particular distribution of one component of overall well-being, whether equal or unequal, is not ‘in itself’ unjust; rather, it is only unjust insofar as, and because, it reflects inequalities in well-being overall.

The primary example of health inequalities that reflect inequalities in overall well-being are health inequalities between socio-economic groups – socio-economic class is one indicator of overall well-being, even more so when combined with health. These inequalities reflect systematic disadvantage across several components of well-being; those who have the worst

---

iii Here, we mean well-being in the very broadest sense, as in how ‘well off’ people are, and it might equally be understood as life prospects or resources or opportunities for well-being.
iv Note some nuances with respect to different degrees of substitutability. At the margins, hardly any good is substitutable (see Hausman [17]).
health also have the lowest income and material security, the lowest level of education, the worst working conditions, and so on. Many of the social causes of ill health, such as poverty, are themselves forms of disadvantage, independently of their effects on health.

Braveman’s view – although she does not make claims about the nature of justice as such – resembles Hausman’s view in that she is concerned with the broader picture, and with how being worse off with respect to health is related to the distribution of other goods [19-21]. She defines unjust inequalities in health as those ‘differences in health (or in important influences on health) that are systematically associated with being socially disadvantaged…, putting those in disadvantaged groups at further disadvantage’[20]. Deaton also emphasises that ‘individual welfare is neither health nor wealth but depends upon both. The [social] gradient [in health] means that people who are deprived in terms of income and wealth are also deprived in terms of morbidity and mortality’[22]. Thus, on such a view, one would be concerned both with how health affects income and how income affects health [23].

The second version of the indirect approach focuses on the nature of the causes of health inequalities. We can call this the ‘unjust social causes’ approach. On this type of view, the fairness of health inequalities is taken to be derivative of the fairness of our social institutions: whether or not health inequalities are unjust depends on the social arrangements in which they come about. If our social institutions and their distributive principles are just, the resulting distribution of health will also be just. The most well developed and comprehensive defence of this view is Norman Daniels’ account [24]. Daniels adopts a (modified) Rawlsian framework of justice, and argues that the Rawlsian principles of justice provide an account of when the distribution of the social determinants of health is fair (see also Daniels et al [25]). On this account, justice in health turns out to be secondary to justice more broadly construed, because what counts as a just distribution of health depends on what counts as a just distribution of social resources more generally.

It is possible to take a view that combines elements of both of these versions of the indirect approach. One example of such a view – or, at least, a view which does not fall clearly into either category – is that of Margaret Whitehead. On Whitehead’s well-known definition, health inequalities are inequitable when they are ‘unnecessary and avoidable, but in addition, are also considered unfair and unjust’[26]. Whitehead is concerned with ‘unjust social or environmental influences’[26], which suggests that health inequalities are unjust when they are the result of causes that are unjust. In this respect, Whitehead’s view resembles Daniels’. And like Daniels’ view, it presupposes an independent account of fairness that allows us to
establish when the relevant social and material circumstances are unjust. Which health inequalities are unjust will ultimately depend on that account’s prescriptions. But Whitehead also indicates that health inequalities that result from the clustering and reinforcing of disadvantage are particularly unjust, and in this respect her view resembles Hausman’s view, in reflecting a concern with systematic disadvantage and inequalities in ‘overall’ well-being.

**MEASURING HEALTH INEQUALITY**

One’s conception of health equity will have implications for what measures of health inequality are most likely to be relevant, which, in turn, will matter for how RCTs are designed, reported, and interpreted. Different health measures yield different kinds of information about health distributions, and different kinds of information may have different degrees of relevance depending on one’s underlying moral concern.

One measurement issue that has been much debated in the health inequality literature is the use of univariate or bivariate measures of health inequality (see e.g. [4, 17, 19, 27-29]). In general, we can think of the bivariate measure as already linked to existing theories of justice (and the indirect framework of equity), whereas the univariate measure is more open-ended (and linked to the direct framework of health equity).

The bivariate measure has usually been favoured by those who see justice in health as interrelated with social justice more generally (see e.g. Braveman et al [19] or Hausman [19]). Bivariate measurements measure health in combination with another group attribute, such as income, education, occupation, race, sex, or rural or urban residence. Thus, the bivariate measure reflects how health outcomes or expectancies vary with other factors that determine or are associated with overall well-being. A bivariate measure is also informative if one is particularly concerned with health inequalities that have social causes, or if one believes that inequalities due to chance are not unfair. Generally, a bivariate measure is more likely to direct us towards systematic causes. Firstly, the effects of chance may be factored out within a group because the bivariate measure gives us the average health of the members of the groups. Secondly, a systematic pattern of inequality between groups is unlikely to be caused by pure chance and suggests a non-arbitrary cause. Thus, Braveman et al argue in favour of the bivariate approach, on the grounds that ‘[e]quity does not refer to all health disparities—for example, in the United States, the average birth weight of girls is lower than that of boys, but this disparity is unlikely to reflect inequity. Equity concerns a special subset of health
disparities that are particularly unfair because they are associated with underlying social characteristics, such as wealth, that systematically put some groups of people at a disadvantage with respect to opportunities to be healthy’ [19].

On the other hand, Christopher Murray argues that health inequalities should be measured comprehensively, rather than selectively, and thus favours the univariate approach. The univariate measure measures only a single variable, the health outcome or expectation of each individual, and as such gives us the ‘pure’ health distribution. Furthermore, the univariate measure reflects the total amount or range of health inequality in a given population – the unit of analysis always remains the individual, and therefore there is only one way in which the health inequality of that population can be represented (for a given health indicator) [27]. In contrast, the bivariate measure shows us health inequality as ‘filtered’ or organised in a certain way, typically social factors like income or social class. Thus, the bivariate measure, unlike the univariate measure, does not reflect the ‘total’ inequality in the population [27]. Furthermore, the bivariate measure often reflects health as a group average, and thus fails to capture within-group inequalities [27, 29, 30]. Murray is concerned that the bivariate approach excludes health inequalities that are not associated with income or class but may nevertheless be unfair, on his view, such as health inequalities caused by inequalities in ‘environmental sanitation, water supply, health services, and social norms about risk factors’ [16].

The univariate approach is interesting, and relevant for RCTs, because the social determinants usually explain only a fraction of the observed inequality. As described above, a direct approach to health equity includes a broader set of health inequality as potentially unfair, not just those reflecting inequality in overall well-being or are associated with socially organised determinants of health [31]. Those who agree with Murray, that all inequalities associated with factors beyond individual control, such as lack of a proper immunization programme or the result of the genetic lottery, may be unfair, may argue that these differences should also be measured in RCTs. The scope of our measures should be broad and include all inequalities in health between individuals (as well as non-health factors, if relevant). Then inequalities can be identified as fair or unfair at a later stage of the analysis [32, 33].

**THE ROLE OF EQUITY FRAMEWORKS FOR EQUITY-RELEVANT RCTs**

So far we have presented an overview of three major health equity frameworks. How are these frameworks linked to conclusions about equity based on the experimental design of
RCTs? A recent paper defines an equity-relevant RCT as an RCT that assesses the effects of an intervention on health equity [9]. Its objectives might include increasing our understanding of, as well as gathering evidence about, how to reduce unfair inequalities or prevent such inequalities from arising. RCTs measure the effects of an intervention. The basic idea of equity-relevant trials is that: a) average differences in primary and secondary outcomes between the control and intervention arms can be attributed to the intervention; and b) that beyond average effects, it may also be possible and desirable to look at impact on distributions within each arm of the study. The design presumes a certain degree of variation in primary and secondary outcomes, and/or in characteristics of individuals or groups that allow for measuring a difference between the arms. Precisely which characteristics should be included depends on the chosen equity framework. If effects of an intervention are unequally distributed, the analysts can evaluate whether this change should be considered an equity improvement or not.

With respect to RCTs, several additional, normative judgements are required to help translate a theoretical understanding of equity into the relevant context. Such judgements concern various practical and methodological issues, which we discuss in further detail below. The results of such considerations provide a complete narrative that helps to clarify which contextualized definition of equity has informed a particular RCT. Furthermore, including an explicit justification of the choice of equity framework will enable others to critically assess the normative validity of the study.vi

FROM GENERAL FRAMEWORKS OF HEALTH EQUITY TO CONTEXTUALIZED STUDIES OF DISTRIBUTIONAL IMPACT

In order to apply a general framework or account of health equity to a specific setting, such as for the purposes of a particular RCT study, the framework must be contextualized, i.e. interpreted and specified in relation to the relevant setting. Researchers conducting RCTs need to base their approaches on what is practically feasible, but may nevertheless translate theoretical work on equity into this ‘real world setting’ [34]. This process of translation requires researchers to make additional normative judgements about how to apply a general

---

vi A ‘contextualized definition’ here denotes an interpretation and application of the general framework to a concrete setting. We discuss this issue further in a later section in the paper.

vi We acknowledge that standardized reporting formats may not allow the space for this, and suggest that this information could be included in an appendix.
concept of equity in a specific context. Examples of such judgements include the delimitation of what should be considered relevant concerns in a particular context, the identification of acceptable levels of inequality versus conclusions about baseline inequity, consistent choice of variables, reporting of findings, choice of measures of inequality, and how to assess whether a change in the distribution of relevant outcomes constitutes an improvement in equity or not. Such judgements are typically not adequately predicted or described within a theory or general definition itself; general theories or definitions usually do not encompass conditions for their application to a specific context or setting. In the following, we will consider, step-by-step, how a health equity framework can be translated into judgments about equitable impacts of health-related interventions (see Table 1 for a summary). We thereby aim to offer guidance for systematic reflection on different aspects of equity-relevance in RCT research throughout planning, conducting, and reporting.

Translating measurements of equality into conclusions about equity

In order to translate quantitative measurements of the distributional impact of intervention outcomes into normative conclusions about impact on equity, an overall normative justification is required – with respect to what kind of equality, and equality in relation to whom, we should consider fair. Independently of the kind of equity-relevant RCTs one is conducting, this overall normative argument can be broken down into different kinds of equity judgements that apply to different stages of the research process. We propose that the different stages of the planning and conducting of an equity-relevant RCT can be distinguished as follows: i) designing a study and choosing equity-relevant outcomes and stratifying factors; ii) data collection; iii) analysing and reporting distributions; iv) drawing conclusions regarding the impact on equity; and v) if appropriate, suggesting recommendations in accordance with the normative conclusions (e.g. implications for further research, or for how the findings can provide input to policy processes). In the interest of keeping it simple, we combine the two first and the two latter stages, respectively, and distinguish between: i) Designing equity-relevant RCTs and collecting equity-relevant data, ii) analysing and reporting health equity-relevant data, and iii) drawing conclusions regarding health equity and recommending further actions.

---

vii As discussed above, recommendations should be limited to implications for further research and outlining how the findings can provide input to policy processes.
Ideally, the justification of contextualized stage-dependent, equity-relevant considerations should cohere with a reasonable theoretical understanding of health equity. That is, the justification should be traceable and formulated in accordance with one of the theoretical accounts of health equity - or at least a consistent framework derived from such an account. All additional normative judgements called for in the translational process should be reported as part of the overall understanding of the particular equity definition of the RCT in question. In the following, we elaborate on the different stages of the research process.

i) Designing equity-relevant RCTs and collecting equity-relevant data

In the first stage of the research process, choices are made with respect to primary and secondary outcomes and stratifying factors that allow for equity assessments. Furthermore, it must be decided whether measuring impact of interventions across individuals or stratified groups (or both) is the best approach.

If the study overall concerns effects on people who are already defined as disadvantaged, a normative argument is needed to explain why this part of the population is disadvantaged and how the perceived disadvantage is related to judgements about health equity. If, instead, the study aims to map differences in effect across groups within a population, a normative justification is required to account for which stratifying factors would be relevant for drawing conclusions regarding health equity. Welch et al [6] helpfully recommend considering the ‘PROGRESS+’ categories as potentially relevant stratifying factors for equity assessments: place of residence, race/ethnicity, occupation, gender, religion, education, socioeconomic status, and social capital—plus age, sexual orientation, and disability [35].

While the former type of RCT makes specific judgements about who should be considered disadvantaged [6], the latter leaves that to some degree open (by identifying who is adversely affected).

ii) Analysing and reporting health equity-relevant data

In the stage for analysing and reporting on health equity-relevant data, new, specifying arguments are required for the complete, normative narrative of particular RCTs. To ensure transparency, decisions made in this stage must be clarified and justification should be offered for the selection of data to be reported. Different perspectives or interpretations of data may
pull in different directions and should be addressed: How should the distribution of effects be presented (e.g. relative versus absolute), and why? If any potentially biasing stipulations have been made with respect to uncertainty in the data set, how are they justified and how does it affect the distribution of effects? How is the cut-off level that distinguishes between equality and inequality justified? All of these concerns come as additional specifications of fundamental theories about health equity, and researchers should aim to report accordingly.

**iii) Drawing conclusions regarding health equity and recommending further actions**

When drawing conclusions about equity with respect to the measured distribution in a given study, a normative argument justifying why the results represent a positive or negative impact on equity is required. How do the results reflect inequity, given the original, general definition structuring the design of the study (for example, with reference to Whitehead’s definition of health equity, how would ‘avoidable’ and ‘unnecessary’ be specified in a given setting)? If the aim is to gather evidence on how to reduce existing inequities, does the equity theory that structures the design of this RCT cohere with the theory that supports the categorisation of identified inequalities as ‘inequities’?

In the case of RCTs, in describing a distribution as equity-relevant, one has conceptually acknowledged that the resulting - or already existing - inequality is unfair. Thus, it is implied that action ought to be taken to eliminate, reduce, or prevent the inequity. This should be spelled out in the report. However, it does not follow that all equity-relevant RCTs need end up with specific policy-recommendation; that could rather be the outcome of broader political decision-making processes.

Table 1: Framework for contextualised equity definitions
| Are you studying the effects (in health- or non-health benefit) in a population considered to be disadvantaged? | **Clarify and contextualize:**  
- What do you mean by 'health equity' in this particular context?  
- Who are the disadvantaged?  
- Why is their disadvantage a matter for equity considerations? | **Clarify and justify:**  
- How is the selected, presented data chosen?  
- What was the considerations with respect to:  
  - what stratifying factors to present?  
  - how to present the distribution of effects?  
  - methods of measuring (in)equality and consequences for what is measured? | **Clarify and justify:**  
- How does the presented data reflect health (in)equity?  
- Is there consistency in how equity has been accounted for throughout the research process?  
- What further actions are required by the findings of this study?  
- Are new studies (quantitative or qualitative) required?  
- How may the data and knowledge generated by the study be relevant for policies, and in what ways can it provide input to the policy process? |
| --- | --- | --- | --- |
| Are you studying differences or equivalence in effects (in health- or non-health benefit) across stratifying factors? | **Clarify and contextualize:**  
- What do you mean by 'health equity' in this particular context?  
- Why are the selected stratifying factors included? | **Clarify and justify:**  
- Why is (in)equality in the distribution of effects (in)equitable?  
- Are there potentially equity relevant biases in the material?  
- Are there any alternative equity-relevant ways of reporting?  
- Are any potentially biasing stipulations made with respect to uncertainty in the data?  
- Where is the cut-off level that distinguishes between equality and inequality in effects put statistically, and why? | **Clarify and justify:**  
- Are new studies (quantitative or qualitative) required?  
- How may the data and knowledge generated by the study be relevant for policies, and in what ways can it provide input to the policy process? |
| Are you studying a gradient of effects (in health- or non-health benefit) across stratifying factors? | **Clarify and contextualize:**  
- What do you mean by 'health equity' in this particular study?  
- Why is the (in)equality in health outcome/access (in)equitable? | **Clarify and justify:**  
- Are there any potential positive or negative side-effects of the intervention? If yes, how | **Clarify and justify:**  
- Are new studies (quantitative or qualitative) required?  
- How may the data and knowledge generated by the study be relevant for policies, and in what ways can it provide input to the policy process? |
| Are you studying inequality (in health- or non-health benefit) across individuals (independently of stratifying factors)? | **Clarify and contextualize:**  
- What do you mean by 'health equity' in this particular study?  
- Why is the (in)equality in health outcome/access (in)equitable? | **Clarify and justify:**  
- Are there any potential positive or negative side-effects of the intervention? If yes, how | **Clarify and justify:**  
- Are new studies (quantitative or qualitative) required?  
- How may the data and knowledge generated by the study be relevant for policies, and in what ways can it provide input to the policy process? |
Table 1: Tentative framework of elements that go into a contextualised definition of equity in RCTs. Valid conclusions on (in)equality require transparency and report on all equity-relevant judgements being made throughout the process of designing, conducting and concluding on the equity relevance of a RCT. The list is not necessarily exhaustive, but can serve as a starting point for further mapping.

| -Why are no selected stratifying factors included? | may they influence (in)equality in health outcome or access? |

Translating theory into practice: Overall concern

Given the various normative assumptions and considerations that may underlie the conclusions drawn about health equity on the basis of an RCT, transparency is called for in all substantive judgments that shape the conclusion about equity impact. Justification of the normative assumptions and considerations can to some extent be provided with reference to a more general health equity framework. In addition, further specifications of normative assessments required for translation to take place should be accounted for. We have here outlined some concerns as a starting point. We believe the experiences of researchers conducting equity-relevant RCTs are called for to provide further detail and to expand on this overview, and will warmly welcome such contributions.

CONCLUSION

In this paper, we have considered how equity-relevant randomized controlled trials can be linked to philosophical theories of health equity. Our purpose was to explore how the evidence base on effective interventions to reduce health inequity can be strengthened by making it more conceptually coherent and transparent. Our approach has been twofold. First, we have provided an overview of three major health equity frameworks and their different implications for the measurement of health inequality. Second, we have described a flexible, step-wise approach to conducting equity-relevant trials that covers a wide range of equity-relevant concerns throughout the various stages of the research process. We do not claim that our account is exhaustive, but rather intend for it to provide a starting point for further discussion as we gain more experience with equity-relevant RCTs.
REFERENCES


