Ethics in action: Approving and improving medical research with human subjects

de Jong, Jean Philippe

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
JUSTICE IN RESEARCH METHODS

Standardization in randomized controlled trials

Jean Philippe de Jong

Abstract

The principle of justice is a crucial aspect in the ethical evaluation of medical research with human subjects. The principle requires that the burdens and benefits of participating in medical research should be distributed fairly among the patient population, but also that the benefits related to the results of medical research should be distributed fairly; the principle is also about whether relevant knowledge is available to aid the treatment of illnesses in different patient groups. In this article I explore the idea that the choice for a specific research methodology can affect how the benefits of research results are distributed within the patient population. Specifically, I suggest that the current ‘gold standard’ methodology for studying interventions, the randomized controlled trial (RCT), can, in particular cases, lead to a distribution of research results that is unfair to disadvantaged patient groups.

I argue that three characteristics of RCTs, all related to standardization, can potentially lead to an unfair distribution of research results: (1) RCTs tend to standardize study groups, and results in standardized study groups can be difficult to extrapolate to disadvantaged groups outside a study; (2) RCTs standardize the delivery of health care within studies, and disadvantaged patient groups outside a study may have more difficulties adapting to these standardized interventions; and (3) RCTs

1 Submitted for publication
require that interventions are standardized, which can mean that less relevant knowledge will be generated on more complex interventions such as lifestyle interventions, which are needed especially by disadvantaged patients.

I conclude that, although standardization in RCTs can help to generate valid knowledge, in particular cases it can also mean that less relevant medical knowledge is available for disadvantaged patient groups. This is an injustice that could be partly remedied by the efforts of good clinicians and by putting less emphasis on standardization in (particular) RCTs. In addition, I suggest a more radical solution to do justice to the multitude of differences between patients: stop considering RCTs to be the gold standard approach for the study of health care interventions, and turn in relevant cases to methodologies in which standardization is a less central concept, such as observational studies and especially qualitative methodologies. These are considerations that could be taken into account in ethical oversight.

6.1 Introduction

Besides the principles of respect for persons and beneficence, the principle of justice is a crucial aspect in the ethical evaluation of medical research with human subjects (Emanuel and Wendler 2000, Kahn et al. 1998). In this article, I argue that the choice for a specific research methodology can affect how the benefits of research results are distributed within the patient population. Specifically, I suggest that the current ‘gold standard’ methodology for the study of health care interventions, the randomized controlled trial (RCT), can in some cases lead to an unfair distribution of relevant knowledge on the treatment of illnesses in different patient groups, affecting disadvantaged patient groups in particular.

Important drivers for the attention to justice in medical research have been scandals involving vulnerable research subjects, such as the Nazi experiments on Jews and prisoners, and the Tuskegee syphilis study on poor African-American men. In the Belmont Report, the question of justice in research is addressed as follows: “Who ought to receive the benefits of research and bear its burdens? This is a question of justice in the sense of “fairness of distribution” (U.S.National Commission for the Protection of
In this article I define an ‘injustice’ or ‘unfairness’ (terms I use interchangeably) as an uneven distribution of relevant knowledge among people with medical needs; in such a way that it especially affects people who are already disadvantaged, for example in terms of the level of education, work and living conditions, income, social position/autonomy, psychological vulnerability, or as a rough measure that encompasses many of these disadvantages: socioeconomic status.

Regulations in many countries currently offer special protection to vulnerable groups such as prisoners, pregnant women and children in order to distribute the burdens of participation in medical research fairly among different patient groups (Emanuel and Grady 2007, McCarthy 1997). In the 1980s, access to research also came to be seen as a question of justice and some of the protections provided by regulations were increasingly perceived as overly protectionist: AIDS activists and other groups with high unmet medical need claimed that they were being treated unfairly because they were denied access to studies and could not benefit from participation in medical research (Emanuel and Grady 2007, Mastroianni and Kahn 2001).

These are examples of concerns about a fair distribution of the burdens and benefits related to participation in medical research. A parallel development has been a concern for a fair distribution of the benefits related to the results of medical research; so, whether research will lead to relevant knowledge to aid the treatment of illnesses in different patient groups (Allmark 2004). To correct for injustices in the distribution of relevant medical knowledge, research on diseases prevalent in specific groups has been stimulated in recent years through grants and regulations. Examples of this are research on rare diseases stimulated by ‘orphan drug’ policies, on diseases in developing countries under the header of ‘neglected diseases’, and on diseases more prevalent in specific groups, for example on diabetes in people with a lower socioeconomic status, and on hypertension in people of African descent. A recent overview of these developments is provided by the WHO Priority Medicines report 2 (Kaplan et al. 2013).

So, to date the concerns for justice in medical research have focused on who should receive the burdens and benefits of participation in research and what diseases should be the topic of research. However, in this article I will explore the idea that the choice of the research methodology is also relevant for justice in research. More
specifically, I will argue that the current ‘gold standard’ for studying health care interventions, the randomized controlled trial (RCT), can, in particular cases, lead to an unfair distribution of relevant medical knowledge. Currently, randomized controlled trials (RCTs) are widely considered to be the ‘gold standard’ of methods for studying health care interventions (Sackett et al. 1996, Timmermans and Berg 2003, p. 27): the bulk of medical research is conducted in the form of RCTs (Getz 2010); RCTs are considered to deliver the highest level of primary evidence for improving health care (Centre for Evidence-based Medicine 2013); and they are of key importance for regulating the market access of medicines (Rägo and Santos 2008).

In the next three sections I will analyze three characteristics of the methodology of RCTs and argue that they can, in particular cases, lead to an unfair distribution of the benefits related to the results of research. In section 2, standard patients, I discuss how control and intervention groups are selected in RCTs. In section 3, standardized care, I discuss how RCTs affect the delivery of care by using protocols. In section 4, standardized interventions, I discuss what kinds of interventions can be studied with RCTs. I illustrate my argument by using examples of research on hypertension management. In section 5, I conclude by suggesting that medical research could lead to a fairer distribution of relevant knowledge if RCTs would not be considered the gold standard and if more studies would use observational, and especially qualitative, methodologies.

6.2 Standard patients

RCTs have a long history of ‘standardizing’ patients, which can lead to an unfair distribution of the benefits related to the results of medical research. I define ‘standardization’ as the process of making something conform to a fixed norm. One of the key aspects of the methodology of RCTs is dividing patients into two groups: an intervention group that is treated with the (new) health care intervention, and a control group that is treated with a control intervention (a placebo or regular care). Besides a difference in the interventions, the study groups need to be as much alike as possible because differences between groups can introduce bias and skew the outcome of a
study. Also to decrease ‘noise’ (random error) and increase the precision of measurements (Hill 1951), RCTs therefore have a tendency to focus on standard patients. In practice, this means that RCTs have often only studied the ‘40 year old white male’ (Epstein 2007), and excluded patients with ‘special’ characteristics (women, children, elderly and ethnic minorities). This practice of excluding these ‘special’ groups also fitted well with the ethical concerns about protecting vulnerable groups from research.

However, to be able to use the results of a study on ‘standard’ patients to guide decision making in regular clinical practice, health care providers needed to extrapolate from middle-aged white men to individual patients, which was not without risk. For example, it could turn out that women or children experienced severe side-effects much more often than the ‘standard’ patient. So, by conducting studies on standard patients, other groups, such as women, children, elderly and ethnic groups, benefited less from the results of research. For instance, although there are many adequately studied treatment options available to treat hypertension in men, there is still a lack of proper knowledge on antihypertensive medications for use with pregnancy-induced hypertension.

To remedy the injustice caused by a focus on standard patients, regulations and policy currently require that a study population is not restricted to middle-aged white men but represents the general patient population. In particular, the inclusion of women, children, elderly, and ethnic groups is encouraged (Epstein 2007, European Commission 2004). For example, for the evaluation of new anti-hypertensive medicines it is now recommended that “patients from relevant demographic subsets should be studied, including both men and women, racial/ethnic groups pertinent to the region, and both young and older patients” (ICH Steering Committee 2000). However, averaging the results of a study with a representative study population might make these results in theory more applicable to the ‘average’ patient, in clinical practice this average patient does not exist: it still requires an extrapolation from the study population to individual patients, with its accompanying risks. This is a problem that especially affects smaller subgroups (e.g. infants, ethnic minorities, pregnant women etc.) because data on them will be ‘swamped’ by those of larger patient groups. For example, it was only discovered after many years of use in clinical practice that beta-
blockers, one of the most frequently prescribed medicines for lowering blood pressure, are not effective in people of African descent (Brewster et al. 2004).

Although extrapolating from an average study population can in principle be problematic for every non-average patient group, I believe that it is especially problematic for disadvantaged patient groups. It is difficult to see how generally well-off groups such as the very rich, or people with a particularly high level of education would respond differently to an intervention (for example by suffering from more severe side-effects), compared to the average patient. This is supported by the fact that these groups do rarely receive special attention in clinical practice. On the other hand, it is a known fact that disadvantaged groups such as children, women, elderly, and ethnic groups can react differently to an intervention.

The acknowledgement of the problems with extrapolation has led to additional regulations and policies that require a separate analysis of subgroups to get insight in any relevant differences (Allmark 2004, Epstein 2007). This has been implemented by designating particular subgroups, defined by sex, age and ethnicity, as the relevant groups to be analyzed. This has led to valuable knowledge on how certain medicines affect particular disadvantaged groups, and in a more equal distribution of the benefits of medical research (Epstein 2007).

However, which patient groups are considered relevant has been standardized: sex, age and ethnicity are currently the standard subgroups. So, again, research subjects have been standardized, only now not according to a single standard but by using multiple standards, a phenomenon we could call ‘niche-standardization’ (Epstein 2007). Although I admit that ‘niche-standardization’ has helped to generate relevant knowledge for some of the disadvantaged groups, I argue that it still might not be doing justice to all disadvantaged patient groups. The current standard subgroups are not based on scientific evidence, but have been shaped by an alignment of classical biological distinctions with long-standing political categories (Epstein 2007) and are used irrespective of the disease or medicine under study. I think that subgroupings should be based on a thorough scientific analysis, taking into account factors such as genetic makeup, physiology and comorbidity, and that it should be decided on a case-by-case basis which ones are relevant. Furthermore, subgroupings based on psychological and sociological distinctions could turn out to be just as relevant for
particular medicines and diseases as biological subgroupings. For example, in the study of anti-hypertensive medication it could make sense to study subgroupings according to socio-economic class, since socio-economic class is strongly correlated to lifestyle factors known to affect hypertension, such as diet, body weight, activity and stress. Yet currently, relevant guidelines make no mention of those factors and focus on the standard biological subgroupings defined by sex, age and ethnicity (European Medicines Agency 2010).

I believe that it is possible to define other, non-standard subgroups and include these in RCTs in order to generate knowledge that is directly relevant to disadvantaged groups and thereby do better justice to these groups. However, setting up criteria for defining these groups, and generating hypotheses about whether they might be relevant in particular cases, would require using other types of methods besides RCTs, for instance observational methods. Furthermore, studying many kinds of subgroups would go against the strive in RCTs to reduce variation that can skew the outcome of a study or decrease the precision of measurements. Moreover, including additional subgroups could mean that a large number of research subjects is required to attain adequate statistical power, with the associated burdens and costs.

In sum, I have argued that RCTs have a long history of standardizing patients, and efforts to do better justice to subgroups of patients tend to (niche-)standardize patients. Because results in standardized study groups can be difficult to extrapolate to disadvantaged groups, this can lead in particular cases to an unfair distribution of relevant medical knowledge.

### 6.3 Standardized care

In this section I argue that the way interventions are used within the context RCTs – as standardized care – can lead to an unfair distribution of the benefits of medical research. RCTs assign patients randomly to either an intervention or a control group in order to ensure that physicians cannot distort an even distribution of patients between these groups. This process of randomization is part of broader efforts in RCTs to prevent variability in the delivery of health care from distorting the outcome of a study (Hill
Variability is minimized by standardizing the delivery of health care within a study, and laying this down in detail in a study protocol (Chalmers 2001). The study protocol can specify, amongst others, dosages, time schedules, administration routes, the characteristics of patients, additional care, medication adherence measures, the health care setting, treatment monitoring, and safety measures. Health care is subsequently to be delivered by strictly following this protocol.

However, it has long been observed that interventions often perform worse in regular clinical practice than in the setting of a study with standardized care, for instance by having a lower efficacy or more serious adverse effects. This problem is known as the efficacy–effectiveness gap (Eichler et al. 2011). One of the causes of the efficacy-effectiveness gap is the fact that patients in trials differ from patients in regular clinical practice with regard to e.g. genetic makeup, physiology, comorbidity, disease burden, stress, physical activity and diet. This issue has been discussed in the previous section. Furthermore, the efficacy-effectiveness gap depends on differences in how health care is delivered in trials and in regular care. Examples of these differences are: more inappropriate or off-label prescribing in regular care, more co-prescribing with an interacting medicine, more continued prescribing to non-responders, more medication errors, poorer adherence to prescribed treatment regimen, more treatment discontinuation, more taking of ‘drug holidays’ and more inadvertent overdosing in regular care (Eichler et al. 2011).

Although this efficacy-effectiveness gap can cause problems for all sorts of patients, I suggest that it may affect disadvantaged patients to a higher degree because the care they receive in clinical practice is likely to be farther removed from the standardized care in trials. There are two possible reasons why this may be so: the first has to do with a patient’s capabilities and the second with proper access to care. Firstly, people with little resources or socially deprived people with little social support structures may in some cases be less capable of adapting their life to an intervention in the exact way it has been studied in an RCT. This idea is supported by the fact that patients who have a low level of education, low self-efficacy expectations, a high degree of life stress, or low level of health literacy experience greater difficulty in adhering to prescribed medicine regimens in general and are more prone to discontinue their medication or take a ‘drug holiday’ (Christensen 2004). In the specific case of
hypertension management, research has shown that problems with adherence in ethnic minorities is a main cause for the fact that antihypertensive medications are less effective in these groups (Holmes et al. 2012).

The second reason why for the disadvantaged regular clinical care can differ so much from standardized care in trial settings, is that people with little resources, those suffering from discriminatory treatment, and other disadvantaged groups can have more difficulties in getting adequate support from the healthcare system (Epstein 2007). For example, disadvantaged patients are more likely to be subject to inappropriate or inadequate prescribing, medication errors and inadvertent overdosing (Epstein 2007). A lack of proper support from the healthcare system makes it harder to deliver the intervention in the exact same way that it was studied in the trial setting, widening the efficacy-effectiveness gap.

To close the efficacy-effectiveness gap so-called ‘pragmatic clinical trials’ have been proposed: trials in which care is as similar as possible to regular clinical care. Pragmatic trials study patients that are similar to the regular patient population; this has been discussed in section 2. Furthermore, pragmatic trials strive to make care as it is delivered within a study more consistent with regular clinical care, i.e. they make care within a trial more variable and less standardized (Tunis SR 2003, Ware and Hamel 2011). Although the internal validity of pragmatic trials can be problematic and although they require larger sample sizes (Ware and Hamel 2011), their results indeed better apply to the average patient, which means they are fairer towards disadvantaged patients. However, following the same line of reasoning as in section 2: although pragmatic trials are fairer to the average patient, their results still do not apply properly to patients who are non-average and who have the most difficulties with adapting to a treatment or with getting adequate support from the healthcare system.

In sum, in this section I have argued that RCTs focus on standardized health care interventions, which can result in an efficacy-effectiveness gap. This gap, which cannot be eliminated completely for all patient groups by using a pragmatic trial approach, may affect disadvantaged patients more, amounting to an unfair distribution of the benefits of research results.
6.4 Standardized interventions

In the previous section I have argued that standardized care (the way in which interventions are used) in RCTs can lead to unfairness. In this section I argue that RCTs are more suited to study the kind of interventions that are, or can be, standardized, and that this can be unfair to disadvantaged patients. In particular, I argue that RCTs are generally more suited for studying medicines, which can often easily be standardized, than for other types of interventions such as lifestyle interventions, which can be difficult to standardize. Not properly studying these types of interventions is unfair to disadvantaged patients because they often experience the most problems living a healthy life. For example: socio-economically disadvantaged groups and ethnic minorities more often have unhealthy lifestyles in terms of cardiovascular risk factors like obesity, physical inactivity and unhealthy diet. So, these patient groups could especially benefit from proper medical research on interventions that target these factors.

RCTs require that interventions are, or can be, standardized by requiring that interventions are of constant quality and reduced to their essential element(s). RCTs require that interventions are of constant quality because this helps to reduce random error (noise) in a study and to arrive at clear results (a clear-cut differences in outcome between groups with different interventions). RCTs also require that interventions are reduced to their essential element(s), because reducing the number of components in an intervention helps to eliminate sources of variability. This means that there is a drive in RCTs towards subdividing interventions into their underlying components and to study these components in isolation.

Besides methodological reasons for requiring standardized interventions, this requirement also has its roots in the (historical) link between RCTs and the regulation of medicines. The development and use of RCTs has from the outset been strongly tied to the need to evaluate medicines and regulate their market access, and information about quality and quality assurance measures have been important considerations for regulatory authorities (Streptomycin in Tuberculosis Trials Committee 1948, The Council of the European Economic Community 1965). Regulatory authorities only allow a medicine to enter the market if it is clearly defined and of constant quality: the
composition of a product should be described in detail and the composition of each specimen should be identical and free of any contaminations.

Due to the historical link between RCTs and medicines, it will come as no surprise that most medicines can meet the requirement for standardization. Medicines are generally industrially manufactured on a large scale, which makes it possible to define their composition in detail, to make products of a constant quality, and to purify products so that they contain just a single active ingredient. However, other types of intervention, such as psychological or social intervention (e.g. psychotherapy, social psychiatric interventions, motivational therapy for substance abuse, lifestyle interventions, occupational health interventions, (nursing) home care and the facilitation of informal care), can have more difficulties in meeting the requirement for standardization.

Here, I take lifestyle interventions for hypertension management as an example. Some lifestyle interventions, such as diet and exercise, are by their nature less clearly defined and of less constant quality because they depend heavily on the activities performed by individual health care providers and patients in daily practice. For instance, although it may seem straightforward to advise a patient to eat a certain amount of vegetables and fruit per day, in practice patients will vary considerably in what kind of vegetables and fruit they eat and how they prepare them (raw, cooked, baked etc.). This variability makes this kind of intervention less suitable to be studied with an RCT because it would be hard to get a clear difference in outcome between groups, and to link the outcome to the intervention. Further standardizing a patient’s diet by prescribing some sort of standard diet would not be an adequate solution: it may work for a short while within a trial but is clearly not a realistic option for daily life. Also an apparently simple advice to walk an hour a day will in practice be taken up in different ways: how far and how fast should a patient walk; does going shopping count or should it be just walking; are breaks allowed and for how long, etc.?

A further problem of studying lifestyle interventions in RCTs is that blinding of health care providers and patients, a means of standardizing interventions, is near impossible. Again, this will make treatment groups more variable and possibly biased, leading to less clear differences in outcomes of groups. Moreover, one of the strengths of lifestyle interventions actually is their variable nature and the fact that they can be
tailor-made based on the skills of the health care provider, the health care setting, the characteristics of a patient, a patient’s preferences and his/her current situation. For example, advice to walk an hour a day may be suitable for someone living next to a nice park, but may be less appropriate for someone living in a bad neighborhood; in such a case, it would be better to give a tailor made advice and suggest going to a gym. Standardizing such interventions in the context of an RCT would make them appear less effective than they in reality are.

A final problem of studying lifestyle interventions in RCTs is that they often cannot be reduced to their essential element(s) because they work by virtue of being part of a holistic approach: they combine a number of elements or consist of several interlinked steps. For example, whether a patient’s diet will contain little saturated fatty acids depends on the overall diet and not just on reducing fatty acids by following an advice to eat low fat cheese, especially if it causes a patient to eat more other high-fat products. A further example related to hypertension management is how to educate and stimulate patients to engage in physical activity (simply giving advice is often not enough). Current approaches propose stepwise learning to increase motivation and suggest to address the complete range of individual and contextual factors that can inhibit or enable physical activity (Stokols 1996, Whitlock et al. 2002). An individual’s socio-economic circumstances, cultural background and living environment are such factors. Splitting up such an intervention by only addressing a single enabling factor or focus on just one step in a learning cycle will make them less effective. So again, the point is that standardizing lifestyle interventions in order to study them with an RCT will make them appear less effective.

In sum, medicines have set the standard for interventions in RCTs: interventions should be clearly defined and of constant quality, and reduced to their essence. I have argued that this is a standard that particular types of interventions, such as lifestyle interventions, can have difficulties to meet, which is unfair to disadvantaged patients because they often experience the most problems living a healthy life. Moreover, I have argued that standardizing such interventions in order to study them with an RCT is also not an adequate solution as it can render them less effective. So, insisting on RCTs as the gold-standard methodology for studying lifestyle interventions
can deprive disadvantaged patient groups of proper knowledge on relevant interventions.

6.5 Discussion

In the three previous sections I discussed how RCTs reduce variability in a study by means of standardization. Standardization aims to prevent a possible distortion of the most important measure in RCTs, the effect of the ‘new’ intervention as compared to the control/old intervention (Chalmers 2001). I argued that RCTs strive to reduce variability between intervention and control groups by standardizing (or ‘niche-standardizing’) patients. Furthermore, I argued that RCTs try to prevent variability in how interventions are used in a study through standardization of care with strict study protocols. Finally, I argued that RCTs require that interventions themselves are standardized: clearly defined, of constant quality and reduced to their essence.

A weakness of my study is that it has an exploratory character and is for a large part based on inferences from ‘circumstantial’ evidence. Especially the link between standardization in research methods and unfairness deserves further study, for instance by conducting in-depth case studies of how individual RCTs have affected particular disadvantaged patient groups. Furthermore, although little direct evidence was available to support the link between RCTs and standardization, this link appears to be somewhat more robust because standardization in RCTs is not an isolated phenomenon, but can be seen as healthcare’s version of how variables are often controlled in scientific experiments: by creating an artificial world, the ‘laboratory’, in which variation is reduced by standardization (Latour 1987). Furthermore, standardization in medical research should also be viewed in light of broader efforts to standardize healthcare (Timmermans and Berg 2003).

The reduction of variability through standardization, as it is done in RCTs, has several important advantages. For instance, it can help to reduce systematic error. This will improve the internal validity of research, meaning that the outcomes of a study can justifiably be ascribed to the effects of the interventions. Reducing variability through standardization also helps to reduce noise (random error). This increases the precision
of measurements and results in clear outcomes of a study. A final benefit of standardization is that it helps make the results of studies suitable for comparison and aggregation, e.g. in meta-analyses.

However, although standardization in RCTs helps generate valid results, this does not necessarily mean that those results are relevant to all patient groups. In the previous three sections I argued that the methodology of RCTs can lead to an unfair distribution of the benefits of research results: (1) research results in standardized study groups can be difficult to extrapolate to disadvantaged groups in clinical practice; (2) the focus on standardized health care can lead to an efficacy-effectiveness gap, which can affect disadvantaged patients more; and (3) the focus on clear and simple interventions can deprive disadvantaged patient groups of proper knowledge on more complex interventions, such as lifestyle interventions.

The first two problems can partially be addressed in clinical practice by health care providers. Given a fair amount of clinical expertise, it is possible to extrapolate from knowledge on standard patients and standardized care to a particular patient in regular clinical practice, and adjust the therapy accordingly. The second problem can additionally be addressed by making sure that disadvantaged patients get appropriate support from the healthcare system and by providing them with complementary care to help them adapt to interventions, for instance by stimulating adherence. However, when there is no adequate evidence available at all concerning a (lifestyle) intervention (the third kind of problem), the health care provider is left solely with his clinical judgment to support a therapeutic approach.

Besides addressing the problems of RCTs in clinical practice, a more direct solution would be to address the problems of the methodology of RCTs themselves. As I described in the previous sections, several developments already focus on reconsidering where, to what extent and how to standardize in RCTs. However, I also pointed out that these solutions can have their own methodological problems and do not always lead to adequately addressing the needs of disadvantaged groups.

An alternative solution to lowering the degree of standardization in RCTs would be to shift the focus of medical research more towards methodologies in which standardization is a less central concept, for instance, by focusing more on observational studies. Observational studies study phenomena in settings closer to real life, as opposed
to the laboratory-like settings of RCTs. To what extent, and in what ways other methodologies can be a proper alternative to RCTs is currently debated with regard to making the marketing authorization of new medicines more adaptive (Eichler et al. 2012). It has been proposed that for some types of medicines it should be possible to grant initial marketing authorization on the basis of smaller RCTs with less statistical power, and use observational methods to generate additional knowledge in a later stage (De Jong et al. 2013, Eichler et al. 2012). Such an approach could lead to more relevant knowledge on a medicine for disadvantaged groups. Although this debate is fueled by concerns about innovation and costs in medicine development and not by a concern for justice, it does signify a shift away from the prevailing position that the RCT is the gold standard.

Furthermore, I suggest that it should be considered on a case-by-case basis whether RCTs will result in relevant knowledge for all patient groups, including the disadvantaged ones, or that other methods might be more appropriate. For example, in the previous section I have argued that for the study of interventions that are complex and difficult to standardize, such as some lifestyle interventions, RCTs are not a suitable methodology. I conclude by drawing attention to a particular kind of observational methodology – qualitative methods – that could complement RCTs and could help to generate more relevant knowledge for disadvantaged groups.

The main rationale for highlighting qualitative methods as a method that could do more justice to disadvantaged patient groups is that they strive for the opposite of standardization: we could call this ‘adaptation’. As discussed, RCTs strive for standardization in order to control and reduce variability. They do so in a laboratory-like setting with the goal of a precise measurement of one component of the object of study. The philosophy of qualitative methods is a mirror image of this: a vital aspect of qualitative methods is that in order to get a comprehensive understanding of the how and why of phenomena, qualitative methods specifically search and describe the variability of real life settings. They do so by adapting methods to both the topic of study and during the study also (Taylor et al. 2010). Qualitative methods are more concerned with generating hypotheses than with testing them and more with revealing new kinds of distinctions than with generating evidence on a distinction provided by the researcher (Pope and Mays 2006). As qualitative methods aim to study variation and
distinctions, I believe that this implies that instead of studying standard patients in standard settings, receiving standardized therapy, qualitative studies can look for evidence on non-standard patients in non-standard settings, receiving non-standardized therapy. Thereby, qualitative studies may result in knowledge that is more relevant to disadvantaged groups. Actually, many qualitative studies focus explicitly on cultural and social differences. Below, I give some examples of how a qualitative approach could relate to patients, care and interventions in medical research, and I indicate what bearing this will have on justice.

Firstly, in order to gain insight into what kind of patient groups are relevant for studying a disease and its treatment in RCTs, one may first have to generate hypotheses about how people can be grouped, and how characteristics of a group can interfere with the effects of an intervention (Epstein 2007). Qualitative methods can help to generate these hypotheses and make them reach beyond the ‘usual suspects’ defined by sex, age and ethnicity. For example, it is known that ethnicity is a determinant of how patients explain and manage hypertension (Morgan 1995). However, this does not automatically mean that belonging to an ethnic group is the actual causal factor in this. For example, a qualitative study showed that how patients manage hypertension might be less related to ethnicity itself than to a history of migration (Beune et al. 2006). This suggests that migrants might be a relevant subgroup to study in RCTs on hypertensive medication.

Furthermore, to close the efficacy-effectiveness gap one should not just make care in trials more lifelike, but one should study the gap itself and how it is related to various patient groups. Qualitative methods could help to understand which groups experience difficulty in adapting to interventions and for studying the underlying mechanisms. For instance, a qualitative study showed that poor medication adherence in Surinamese and Ghanaian patients was related to patient worries about negative effects on sexual performance and that they discontinued medication when visiting their homelands (Beune et al. 2008). This knowledge could be used in clinical practice to tailor information to these groups in order to improve adherence.

Finally, qualitative methods appear to be a viable alternative for studying the more complex psychological and social interventions, such as some lifestyle interventions, that are especially needed by disadvantaged patient groups. Qualitative methods can help to study the main strength of these interventions: how several
components are combined into a successful therapy (Mol 2006). Furthermore, they can be used to study how health care professionals exercise clinical judgment and tailor therapies to patients (Lutfey 2005, Mol et al. 2010). For example, helping patients to engage in physical activity requires insight into what relevant enablers or barriers are. For instance, a qualitative study showed that specific barriers for Ghanaian and Surinamese patients are preferences for large body sizes and unfamiliarity with recommended physical activities such as cycling (Beune et al. 2010). Interventions could be adapted accordingly: focus less on weight loss and more on physical activity, and promote walking instead of cycling.

In sum, RCTs tend to standardize patients, care and interventions, meaning that in some cases less relevant medical knowledge will become available for disadvantaged groups. I do not conclude that RCTs should not be used for studying health care interventions, but suggest that RCTs should be considered less of a gold standard. To put the current status of RCTs in perspective, I quote what one of the major contributors to the methodology of clinical trials, Bradford Hill, said in 1952: “The statistically guided therapeutic trial is not the only means of investigation and experiment, nor indeed is it invariably the best way of advancing knowledge of therapeutics. I commend it to you as one way” (Hill 1952). I believe that the benefits that derive from research results could be more fairly distributed if RCTs would standardize less and would be complemented more often by observational, and especially qualitative, methods. This could help to generate knowledge that is more relevant for disadvantaged patient groups and so to better do justice to the multitude of differences between patients. These are considerations that could be taken into account in ethical oversight.


