

Reconceptualising risk–benefit analyses: the case of HIV cure research

Robert Steel 

Clinical Center Department of Bioethics, National Institutes of Health Clinical Center, Bethesda, MD 20892, USA

Correspondence to

Dr Robert Steel, Clinical Center Department of Bioethics, National Institutes of Health, Bethesda, MD 20892, USA; pirateofthecaribbean@gmail.com

Received 30 April 2019

Revised 25 July 2019

Accepted 8 October 2019

Published Online First

15 November 2019

ABSTRACT

Modern antiretroviral therapies (ART) are capable of suppressing HIV in the bloodstream to undetectable levels. Nonetheless, people living with HIV must maintain lifelong adherence to ART to avoid the re-emergence of the infection. So despite the existence and efficacy of ART, there is still substantial interest in development of a cure. But HIV cure trials can be risky, their success is as of yet unlikely, and the medical gain of being cured is limited against a baseline of ART access. The medical prospect associated with participation in cure research thus look poor. Are the risks and burdens that HIV cure research places on participants so high that it is unethical, at present, to conduct it? In this paper, I answer 'no'. I start my argument by describing a foundational way of thinking about the ethical justification for regulatory limits on research risk; I then apply this way of thinking to HIV cure trials. In offering this analysis, I confine my attention to studies enrolling competent adults and I also do not consider risks research may pose to third parties or society. Rather, my concern is to engage with the thought that some trials are so risky that performing them is an ethically unacceptable way to treat the participants themselves. I reject this thought and instead argue that there is no level of risk, no matter how high, that inherently mistreats a participant.

INTRODUCTION

Modern antiretroviral therapies (ART) suppress HIV in the bloodstream to undetectable levels; maintaining an undetectable viral load prevents sexual transmission and lifespans for individuals who maintain viral suppression are very good. Nonetheless, because HIV still persists in reservoirs throughout the body, people living with HIV (PLWHIV) must maintain lifelong adherence to ART to avoid the re-emergence of the infection. And lifelong provision of ART to the near-40 million PLWHIV worldwide is a serious financial and logistical challenge. The development of a cure for HIV is thus of potentially high social value, even given the fact that effective treatment already exists.

Nonetheless, HIV cure trials can be risky, their success is as of yet unlikely and the medical gain of being cured is limited against a baseline of ART access. So the medical prospect associated with participation in cure research looks poor. Are the burdens and risks that HIV cure research places on participants so high that it is *ipso facto* unethically abusive to conduct it?¹ In this paper, I answer with a clear 'no'.

¹This question was raised by a recent special issue of *Journal of Medical Ethics*; see especially the issue introduction.²¹

I start by describing a foundational way of thinking about the ethical justification for regulatory limits on research risk. I then apply this way of thinking to HIV cure trials. In the end, I argue that even studies with extremely unfavourable risk profiles can still treat participants in an ethically acceptable way, provided that both the goals of the study and the methods of subject selection are sufficiently sensitive to the values of those who enrol.

My analysis is confined to studies enrolling competent adults; I do not consider minors or adults who lack decision-making capacity. I also do not consider risks to discrete individuals outside of research who may be harmed by the conduct of a study, nor do I consider potential broader consequences, including, for instance, the possibility that conducting risky research might expose an institution to legal liability or lead to a loss of social trust. My conclusion is compatible with the thought that high-risk HIV cure research might be ethically problematic for such reasons. Rather, my concern is to engage specifically with the thought that some trials are so risky that performing them is an ethically unacceptable way to treat *the participants themselves*. I reject this thought. I argue that there is no level of risk, no matter how high, that inherently mistreats a participant.

PATERNALISM, AUTONOMY AND RISK

This paper builds on the work of Miller and Wertheimer's, which, as I see it, offers two general insights with respect to our specific present topic.¹ The first is that regulatory limits on research risk are *prima facie* paternalistic; the second is that they can nonetheless be legitimated by trade-offs made at the policy level.

Some preliminaries say that an agent or policy acts *paternalistically* when that agent or policy restricts another agent's liberty out of concern for that agent's good. Hence, whether a policy is paternalistic depends not just on the existence of a restriction but on the justification under which it is proposed. Not all justifications for limiting research risks are paternalistic, for instance, it is sometimes argued that such limits are necessary for maintaining social trust. But to the extent that restrictions are motivated directly by a desire to protect participants from unwise choices that would compromise their welfare, they do appear paternalistic. This is troubling given that we stipulated that we are discussing competent adults and competent adults have the presumptive right to make their own decisions.

We can understand that right as grounded in *autonomy*. Autonomy, broadly speaking, is the



© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Steel R. *J Med Ethics* 2020;**46**:212–219.

capacity to choose for oneself among one's options: it involves the ability to process relevant information about the possible outcomes of one's choices and the degree to which they advance or counter one's interests and values. Ordinary adults are presumed to have that capacity and we respect it by affording them the presumptive right to issue authoritative decisions with respect to those matters that are up to them. It is against that right that paternalism offends.

Still, not all paternalism is created equal, and it is useful to distinguish between *hard* and *soft* paternalism. Hard paternalism involves not only restricting an agent's liberty, but doing so in a way that prevents them from making a genuinely autonomous choice. Soft paternalism, by contrast, involves restricting merely non-autonomous choices. Paradigmatically, such choices include ones made on the basis of false beliefs about their consequences—for example, restraining me when I attempt to drink a glass of what I believe to be water, but you know to be acid. If we care about people's abilities to govern their lives according to their own interests and values, then that gives us strong reason to reject hard paternalism, but, by contrast, it may actually give us reason to *promote* soft paternalism. After all, I will have trouble governing my life according to my interests and values once I die from accidentally ingesting acid.

When we transition to the level of policy, attempts to put into place regulations that protect people from unwitting, non-autonomous choices may, as a collateral effect, also restrict others' freedom to choose autonomously. With respect to our present topic, biomedical research is complicated and difficult to understand. Worse yet, potential research participants may be made vulnerable by their illnesses and may be liable to suffer a variety of decisional defects, for instance, due to desperation or therapeutic misconception. The chance that participants may end up making non-autonomous, harmful choices may thus justify putting into place limits on the sorts of trials which people may be offered. For instance, it may justify forbidding trials with a too-unfavourable balance of risks against benefits. But such a prohibition, once installed, limits everyone. It limits even those who fully understood, rationally deliberated, and were ready to autonomously enrol in a riskier study.

Miller and Wertheimer's thought, which I accept, is that the latter limitation could be in itself undesirable while still on balance being justified by its necessary relation to the former. The cost of restraining some genuinely informed and autonomous actors could be more than redeemed by the much greater benefit of protecting *most* people from the *very bad* consequences of the non-autonomous choices they would otherwise make.

I find this justificatory strategy appealing by virtue of its ability to reconcile two thoughts: first, that research participants are competent adults who have the presumptive right to make their own decisions; and second, that there is nonetheless an important role for the regulation of research risk to play in protecting participant welfare. Not everyone will be equally impressed by the *need* to reconcile these thoughts. More robust paternalists may not feel much need to justify pervasive regulation of research risks, and some libertarian types may think that all such regulation should simply be done away with. Against such positions, I have no further argument here. As a result, my conclusions can be read as conditional on accepting this approach.ⁱⁱ

ⁱⁱWhat are some other ethical approaches to the regulation of research risk? Miller and Weijer^{22 23} and London^{24 25} have defended fairly restrictive 'equipose' views. They connect this condition to the duty of care held by physicians and/or to

Still, even accepting this approach, there is much left to be worked out. First, *what* level of policy are we describing, and *who* is the actor instituting the restraints? I focus on Institutional Review Boards (IRBs). When I describe strategies that studies might use to minimise the risk of non-autonomous enrolments, I intend that the inclusion of such strategies in research protocols could be demanded by IRBs as a condition of approval. Furthermore, IRBs doing so can be understood as executing their legal mandate to evaluate whether the risks to subjects involved in a study are reasonable in relation to the benefits.² I suggest that what makes some risks reasonable is that there is a sufficiently low likelihood of their being assumed non-autonomously.

In evaluating protocols, I suggest IRBs ask themselves two questions:

(1) In light of everything involved in this protocol, *what kind of life situation* a person would need to be in, and *what kind of values* would a person need to have, in order to give genuinely autonomous consent to enrol? Who does participation *make sense* for?

And,

(2) Given the recruitment procedures the study plans to employ, how successful do we expect it to be at enrolling (all and only) the people described in the first step?

If the answer to the second question is that those procedures would be successful enough, then the risks the study imposes are not unreasonable with respect to the treatment of its human subjects.ⁱⁱⁱ

Two points of clarification. First, answering these questions requires that we think about who could or would autonomously enrol in a given study, and in some cases, this will require applying a more detailed understanding of autonomy than I have so far offered. I described autonomy as consisting of abilities, like the ability to process information: in reality, these abilities come in degrees and it is not immediately obvious how to relate dichotomous talk of autonomous and non-autonomous agents or choices

oversight duties of the state, and importantly, they think this protection cannot be waived even by competent adults. What equipose-style views imply for the phase I research that serves as the principle focus of this article is not obvious, as some defenders of equipose hold that it is only properly applied to robustly therapeutic research.²⁶ They, at least, may find much in the paper to agree with. In another theoretical vein, Miller and Joffe have argued that participants' assumption of risk should be limited to a maximum magnitude implicit in social morality, obtained by drawing comparisons to similarly situated risky activities—although they do not bother justifying why that social morality, even once discovered, should be enforced on an individual who may not share it.²⁷ Rid and Wendler also do not feel obligated to explain what justifies imposing an upper limit: in developing their framework they find it sufficient to note that 'no morally serious person' would allow killing in research, despite the fact that any limit's foundational justification 'remains an open question'.²⁸ I believe that my approach ultimately has the resources to give more satisfying explanations for why different levels of risk are acceptable in different contexts than existing alternatives. This paper aims to show by doing, that is, to develop some of those explanations; if it succeeds, then that success counts in its favour. Still, the intent is not to directly critically assess, let alone refute, alternatives. Hence, the conditional form of the claim in the main text.

ⁱⁱⁱThe first question is similar to Alex Rajczi's "agreement principle," which also has similar motivations. However, I depart from Rajczi in requiring that the first question is to be asked always in conjunction with the second, and hold that only once the second is answered can we determine the acceptability of a protocol. In effect, Rajczi treats risk-benefit assessments and reliability of the recruitment and consent processes as in principle separable, which I deny.²⁹

to graded abilities. Similarly, in many cases, agents will not be motivated by a single consideration, but by many, and these may correspond to some mix of accurate and inaccurate information. Finally, it is also worth noting that some believe that even values perfectly applied under perfect information need not result in an autonomous choice, if those values are themselves sufficiently defective in their content or origin, as, for instance, when they derive from social conditioning or internalised stigma.

Rather than presuppose a particular theory of autonomy, I attempt to limit my conclusions to those that do not depend on contentious underlying details. Recall that we introduced the notion as a capacity to make one's own choices which is presumptively had by ordinary adults. This is part of the point of a theory of autonomy: to describe a widespread capacity that grounds also-widespread equality and non-interference rights. But ordinary adults routinely make decisions under less than perfect information, with motives that vary in how well they match the truth of their situation, and under the influence of social conditioning and internalised stigma. These include life decisions we strongly respect: marrying, having children, buying a house that pursuing a career. Whatever our ultimate theory of autonomy says, it should not be so demanding that it loses the ability to explain why actual (inevitably imperfect) decisions should be respected. On these grounds, my working assumption will be just that determinations of autonomy, however properly understood, will still involve substantial deference to the expressed values of the agent in question.

Second, this process ends with an assessment of whether the recruitment process will be successful enough at recruiting all and only those people for whom it makes sense to enrol. Two types of recruitment errors will be salient. Type 1 errors occur when lax enrolment standards lead to participants being enrolled non-autonomously; type 2 errors occur when barriers to enrolment end up excluding participants who would otherwise have autonomously enrolled. Both are failures, but it is the type 1 errors that will be my primary focus. I mention type 2 errors only because the goal of minimising each stands in natural tension with minimising the other and because that relationship may be a relevant factor in deciding whether a given rate of type 1 errors it produces is 'sufficiently low'.

How low is 'sufficiently low?' The only way to have a *zero* chance of type 1 error would be to never enrol anyone at all. But that is not a reasonable proposal. Instead, we need some theory of what makes a rate tolerable, particularly in relation to other normatively relevant facts about the research.

I do not attempt that theory here. Instead, I confine myself to making use of the following partial but extremely plausible claim: as personal risks of participation become more serious, type 1 errors become less tolerable and similarly with the reverse. This relative statement falls short of telling us about absolute values, so to arrive at any absolute claims, I will have to additionally appeal to intuition and judgement. In what follows that is what I do.

SHORT, WELL-MONITORED TREATMENT INTERRUPTIONS

As noted earlier, current ART suppress HIV in the bloodstream to undetectable levels, but the virus continues to persist in reservoirs throughout the body. Dealing with the viral reservoir is necessary for a cure, but the mechanisms of the reservoir's persistence and reinitiation of infection are not fully understood. One widely used approach to assessing an intervention's effect on the reservoir in HIV cure research is the *analytic treatment interruption* (ATI). In an ATI, researchers discontinue ART and

chart the subsequent course of the infection—with the most ambitious hope being that the virus does not return at all, but regardless with the aim of collecting data about when and how it rebounds. ART is then restarted.

Because they are presently the gold standard for measuring the efficacy of curative interventions for HIV, ATIs are a common element in otherwise very different studies. What is more, many cure studies will include a control arm, so we can anticipate that there will be participants who receive ATIs in the absence of any other intervention. Indeed, some whole studies use ATIs in isolation from any curative intervention at all. Start, then, by considering the prospects of a participant who undergoes *only* an ATI.

Begin with what we know of ATIs.^{iv} Treatment interruptions were initially investigated in the hope that they would be of use in clinical care. Unfortunately, a large, early trial demonstrated that long, repeated interruptions significantly increased the risk of serious adverse events up to and including death.^v But the shorter, better-monitored interruptions used in contemporary ATI trials do not present the same risks.^{vi} A recent consensus statement underlines a general optimism about the ability to conduct ATIs within acceptable safety limits⁴; indeed, in this context, it is worth noting that one recent meta-analysis has even advocated a reconsideration of using newer, safer ATIs in clinical care.^{vii} I will not enumerate specific risks here, although they certainly exist, but allow that they are mild.

Still, 'mild' risk is not 'zero' risk. Suppose that a person really stood to gain nothing therapeutically from undergoing an ATI. What, then, could motivate them to volunteer for one?

Research into participant motivations reveals that any given individual is likely to have multiple reasons for enrolling. In studies without the prospect of therapeutic benefit, two common motivations, which are not mutually exclusive, are interest in payment for participation and the altruistic impulse to help people by enabling socially valuable research.^{viii} So it is not hard to see what might motivate a person to volunteer for an ATI, the same thing that motivates people generally. For an ATI study with low risks, decent pay, and the prospect of contributing social value, I can see no special reason to believe it would be challenging to recruit people suitably so motivated. Or, in other words, I see no reason to believe special measures would be necessary to maintain a tolerably low type 1 error rate.

That is a fairly bland verdict: a low-risk study with social benefit is *prima facie* approvable. Few say otherwise. But we reach this verdict in an atypical way, and it is worth exploring the ways that our reasoning departs from more traditional risk-benefit assessments.^{ix}

^{iv}For an extremely useful overview, see Wen, *et al*³⁰

^vSpecifically: the SMART study scheduled visits for months 1, 2, and then every 2 months in the first year and every 4 months in the second; in the interruption arm, it demonstrated an increase from 1.3 to 3.3 instances of opportunistic disease or death per hundred person years.

^{vi}A secondary analysis of SMART data using the inclusion criteria of an ATI trial found that for that subgroup, there was no increase of clinical events within the first 16 weeks.³¹

^{vii}The recommendation is targeted to paediatric and adolescent populations, where toxicity and non-adherence are serious concerns.

^{viii}Given their limited prospect for medical benefit and their strong baseline medical prospects, cure research participants are most specifically like healthy volunteers. For research on healthy volunteer motivations, see Stunkel and Grady³², Grady *et al*³³

^{ix}I identify the traditional view of risk-benefit assessments, against which I contrast my own, with Miller's four-step process,³⁴

So bracket, for a moment, the big two motivations of altruism and payment. Instead consider one specific to ATI trials: the motivation to stop taking medication itself, in order to obtain relief from an upsetting routine or out of curiosity about the workings of one's body. Henderson *et al* have reported these motivations among participants in the SEARCH 019 trial in Thailand, and offer those reports as revealing important, previously overlooked benefits.⁶ Eyal, in response, allows that such benefits are genuine, but complains of their triviality—taking a pill a day is 'relatively undemanding', and everyone will eventually be back on ART anyway.⁷ ^x When thinking about the ethics of ATI trials, what are we to make of this conflict in the assessment?

Intuitively speaking, I do not find either pill fatigue or curiosity about one's body to be obviously inadequate to rationalise participation in an ATI study. True, most people who undergo an ATI will quickly rebound, and this must be made clear—but perhaps a short break is worth having. What is more, there is a chance at a longer remission. The largest study to date on post-treatment control found that 4% of the people who began treatment in the chronic phase of HIV infection and 13% of people who began treatment in the early phase later met their definition of post-treatment control. Of those, perhaps something like one in five will maintain control for ≥ 5 years.⁸ The chance for a 5-year break in the pill routine is small but meaningful.^{xi}

Being limited to an intuitive assessment of whether this choice could be autonomously made is a defect, and it is a product of the earlier decision not to settle on any particular theory of autonomy. Nonetheless, it is worth noting that autonomy is in one sense a *less demanding* standard than objective correctness: perhaps few options are objectively correct to choose, but many more can be autonomously chosen. This is part of what can ground reasonable confidence that a person could autonomously choose to enter into an ATI trial, even without having to suppose that it presents an *actually good deal*, either medically or overall.

This reveals the first significant novelty in our reasoning: traditionally, risk–benefit assessments present themselves as discerning the degree to which participation is, or is not, objectively beneficial (and then asking, if it is objectively harmful, whether offering it is nonetheless justified by the social value of the research). We instead address the question of whether it makes sense given a particular person's particular values and priorities. That question *can* be easier to answer. For example, when a person has strong but idiosyncratic desires (here, perhaps, in valuing a pill holiday) or when a person desires to make personal sacrifice for others (as with extreme altruists), it is often very hard to tell whether they are made objectively better or worse off by getting what they want. But it can be easier to determine that their pursuit of it is fully intelligible.^{xii}

which London refers to as 'the common rule approach'.²⁴ I take it to be the most fidelitous interpretation of the classic synthesis due to Emanuel *et al*.³⁵

^xTo be clear, Eyal does not oppose the trial or say that it has an inadequate risk–benefit profile. What he objects to is what he sees as Henderson *et al* reclassifying a (potentially permissible) bad prospect trial as a good prospect trial via the overemphasis of marginal benefits.

^{xi}It is worth noting that even absent daily medication, people on breaks will have to submit to regular testing. A medication break is not a full break from medical involvement.

^{xii}This is not to say that it will *always* be easier to determine that a choice is autonomously made than to determine an objective risk–benefit profile—just that it *can* be in some important cases, including the highly salient one where we have to find some way of weighing altruistic motivations.

Given all that, I say we should be willing to furnish a positive answer: there are some people for whom it would make sense to enrol in an ATI trial just on the basis of a desire for a break from medication. But that just answers the first question. We still have to consider how successful we think a study could be in enrolling *only* those people.

Here it is useful to consider the results of another study from Henderson *et al*, this one assessing levels of decisional conflict among a small group of participants and non-participants in an HIV cure research study. Overall, she finds that both groups scored highly on a measure incorporating self-assessed informational clarity, support in decision-making and satisfaction with a choice. She also found that all eight of those who did participate in the research were initially motivated in part by the prospect of stopping ART, all eight reported elements of the treatment interruption as positive and all eight remained satisfied with their initial choice.⁹ ^{xiii}

Retrospective satisfaction with a decision is not a definitional of it being autonomously made, but it is evidence counting in favour. Henderson *et al*'s results leave me cautiously optimistic not just that a hypothetical person *could* autonomously enrol in a study seeking a break from medication, but that such people are identifiable enough to be reliably recruited. Granted, this is a single study with a small sample. Nonetheless, it is suggestive of the possibility of designing studies which would properly recruit subjects on the basis of their desire for a medication break, and so suggestive that a tolerably low rate of type 1 error could be achieved even without—as we earlier did—appealing to either payment or altruism.

That conclusion exposes the second novel feature of this reasoning as against more traditional views. Here the relationship between social value and the justification of individual risks is not one of necessity. Yes, it is often true that participants will be motivated to enrol by the opportunity to facilitate social goods, and if that is a dominant motivation they act under, then in order for their enrolment to be autonomous, the study better stand to produce those goods; a social value requirement is indirectly secured. But the desire to produce such goods is only one motivation suitable to explain autonomous enrolments. When other motivations are available, the study needs only genuinely match *those*.

Still, the practical impact of this difference is limited. First, despite the fact that this new structure implies that a suitably low rate of type 1 error might be achieved *even by* a study lacking in social value, there are independent reasons to insist on social value (a point we return to in section 5).¹⁰ Second, even in this case, all of the participants surveyed unsurprisingly *also* reported altruistic motivations in addition to their desires for a treatment interruption; as noted, people have mixed motivations. Whether any participants would be interested in a study that *only* offered a treatment interruption without the altruistic upshot is unclear.

So despite its important differences from more traditional views, the theoretical perspective here still yields a pretty standard result. This changes, though, when we move to consider higher-risk studies.

^{xiii}This is not to say that the picture was entirely rosey: two of the eight reported that though satisfied with their choice, they might not make it again. Those two seroconverted during the interruption, which leads from testing negative to testing positive on HIV tests. Participant understanding of this risk was varied and this experience suggests that it should be emphasised. Perhaps seronegative status should be an exclusion criterion where feasible. Although I classified the risks of short well-monitored ATIs as overall 'mild', I think that seroconversion, particularly in jurisdictions where for example, employment may be jeopardised, is serious and should be treated as such.

LESS SAFE ATIS, TOXICITIES AND SIGNIFICANT RISK

Previously we discussed short and well-monitored ATIs with conservative restart criteria. Suppose we instead imagine an ATI with less conservative restart criteria. Suppose investigators wish to allow a period of significant viraemia in order to test whether immune control could eventually re-establish itself; there are reasons to think this is a plausible strategy and some such trials are actively ongoing.^{xiv} The propriety of such designs is thus practically important. On a more purely theoretical level, we could imagine the existence of scientific reasons to test interruptions even more extreme; in the limit, going so far as to tolerate the appearance of AIDS-related events and only restarting ART once they cannot be clinically managed in any other way, though, again, this is a mere theoretical possibility.

In addition to the risks attendant to less conservative ATIs, potentially curative interventions can themselves be risky. Some have serious toxicities. For instance, histone deacetylase inhibitors have been used in an attempt to reverse the latency of the HIV reservoir; however, they are mutagenic and can have serious cardiac, haematological and gastrointestinal toxicities. And inhibitors of checkpoint blockers have been investigated as potential catalysts of a reinvigorated host immune response; however, they may also cause autoimmune reactions leading to diabetes or heart failure.¹¹

It is also worth noting that the psychosocial risks involved in participation can be substantial. In the USA, concerns about discrimination are tempered by a legal regime in which health information is protected and the ADA prohibits many forms of discrimination on the basis of HIV status.¹² But this is not uniform across jurisdictions.

Risks always ought to be minimised insofar as possible. But suppose that we are considering a hypothetical study where, given the scientific aims, it is impossible to avoid a remaining combination of serious risks. Perhaps the cure design requires significant, prolonged viraemia; perhaps the only appropriate cohort resides in a hostile jurisdiction. How would that change our answers to the question of whether there are enrollable participants for whom enrolment *makes sense*, given their values and life situations?

Here's the first difference: it now strikes me as unlikely that there are many people out there for whom *just* an interest in a pill holiday would be sufficient reason to enrol. This is not to pass judgement on the objective importance of various risks and outcomes, which is again not our project. Rather, it is just to state the plausible descriptive claim that few people do, in fact, value pill holidays highly enough.

What, then, of other motivations, motivations which may more plausibly explain an autonomous decision to assume *significant* personal risk? One obvious candidate is again payment. There are plenty of people for whom participation in even very risky research would 'make sense', if it paid enough. However, despite believing the desire for payment to be a fully legitimate motive, I will not appeal to it here. I do this so as to avoid addressing the variety of ethical objections that have been raised to using payment as an incentive to undertake serious risks.^{xv}

Bracketing payment leaves us with the other obvious option: altruism. When we ask whether there is anyone for whom it

could make sense to participate in a risky HIV Cure trial, the answer is yes. It could make sense for those who value the social benefits of research highly enough to shoulder the personal risks of participation. Are such people merely hypothetical? Not apparently. In the HIV community, we have a positive reason to believe that such people exist: consider, for instance, one recent survey of PLWHIV's attitudes toward participation in cure research finding that 34% were willing to accept severe or extremely severe side effects with no prospect of personal benefit.¹³

Granted, a report of willingness to participate is not equivalent to an autonomous choice to enrol, nor is it a demonstration that the person is a good candidate to enrol if solicited. Nonetheless, it is indicative of an attitude. And even if the large majority of those survey respondents would not, in the end, be good candidates, there would still be the fraction that remained. Even a small fraction of a large population, if it can be reliably identified, can furnish the small numbers required for early phase research.

Consider also some direct quotes from participants in HIV cure research. Here are two from an American context¹⁴:

It was so devastating when this happened to me... But now I think there's a reason. I'm supposed to do these studies. In a sense it's kind of a blessing... I figure if I'm going to have this then I'm going to do whatever I can to help someone, the next generation or the one after that. It would make something good out of something bad.

And here are two from a Thai context⁶:

We do not have many chances to be useful in the world... We are guinea pigs, that's right.... The guinea pig is useful... The main benefit may not happen to us directly but it will happen to other people definitely. Even if there is not yet any conclusion, we are in the process. It's like we are on a conveyor belt—when the packaging comes out, if it cannot be used and is thrown away, that package can provide guidance for a new package in the future.

These participants speak movingly of a desire to help others; for a person with these values, enrolling in even substantially risky research with no prospect of benefit nonetheless makes sense as a fully autonomous choice.

But once we commit to altruism as the dominant rationalising motivation, we become subject to corresponding constraints. The weakest is that *if* altruism is the dominant motivation participants act under, then in order for their enrolment to be genuinely autonomous the study must *actually have social value*. Trials which, for example, are not scientifically valid or which never complete enrolment lack social value altogether, and *ipso facto* lacks whatever social value the altruist seeks to promote.

But less trivially, a trial must not just have some social value, it must have the right *kind* of social value given the participants it enrolls. Perhaps there are objective facts about which outcomes are more or less valuable, facts independent of anyone's opinion. Regardless, in pluralistic societies, individuals will have their own views, and we must acknowledge that a person who is willing to make large personal sacrifices to promote social value under one conception may not be willing to do so under another.

Imagine a candidate cure that could lead to cost-effective improvements in health in high-income countries, but which would be unlikely to ever scale to low and middle-income countries. In light of the potential benefits, it would not be unreasonable for researchers to pursue a trial testing it, for an IRB to

^{xiv} These 'set point' designs are inspired by observation of this pattern in primates.³⁶ They are also attractive in light of the observation that some posttreatment controllers experience a similar spike before regaining control.⁸

^{xv} For an argument in favour of payment in this context., see Largent³⁷

approve it, nor for someone who valued that outcome to enrol. However, it would also be reasonable for someone *not* to enrol. Consider a person from a low-income country who was indeed willing to assume substantial risks, if doing so could be of benefit to her co-nationals. If she were to enrol under the (false) impression that this study promoted her altruistic goals, then her enrolment would be less than fully autonomous; it would be a type 1 error.^{xvi}

For those altruistic participants for whom the study's social value is their essential motivation, their ability to autonomously enrol requires that the study's actual social value not depart too significantly from the value they seek to promote. Given that we have been entertaining the prospect that risky studies are permissible precisely *because* they can recruit altruists acting fully autonomously, this means risky studies must be particularly attentive to the possibility of such a mismatch.

Saying that mismatches ought to be prevented is one thing; saying how is another. One obvious answer is to bring the details of a study's potential social value into the consent process, such that participants are informed about things like the likelihood of a cure strategy scaling to low-income contexts. Once so informed, they can then exercise their prerogative to refuse enrolment if they like. This would have the point of consistency in its favour, insofar as it would involve treating the social value of a study in a parallel way to its medical risks and benefits. After all, when it comes to medical risks and benefits, consent forms do not merely state that studies have some unspecified medical risks and benefits. They state, to the best of their ability, *which particular* risks and benefits studies have so that participants can use their own values as a guide to how tolerable the balance between them is.

Unfortunately, there is a reason to doubt the effectiveness of the consent process already, with respect to patient comprehension. Adding more information may make things worse. And given how speculative estimates of future social value are, we ourselves may not know what descriptions are appropriate.

These concerns are valid, but not fatal. Even if we are deeply uncertain of the social value of a study, communicating *that uncertainty itself* may already enhance the ability of participants to make autonomous decisions. Take a study attempting to develop biomarkers for the HIV reservoir, but not promising any immediate therapeutic advances. The results might be significant, and eventually important to advances in care, but they also may never translate to anything at all. If they do, they are likely to do so only on a long time frame. That is something that itself could be said. Perhaps a sample timeline could be given based on a similar past study that did eventually succeed in translation.

It is worth emphasising exactly *how cursory* descriptions of social value are in current consent forms. A boilerplate sentence or two is typical, shared across wildly different studies despite the fact that they in fact aim at very different potential scientific and health benefits. Even a brief and digestible—but study-specific—description would be an advance over the status quo.

It is also worth emphasising that the consent process is just one part of the overall recruiting process. We might also intervene elsewhere before patients even engage in a consent conversation. For instance, researchers could consult with HIV

activist groups to identify promising populations to target for recruitment. Activist groups are likely to be able to find people who are not only altruistically committed to fighting HIV, but who are also knowledgeable and who have access to independent resources to help them understand studies. With a carefully recruited population of prospective patients, an enhanced consent process might be more effective. Indeed, researchers might partner with activist groups to produce digestible materials specifically aimed at participant education for risky but worthwhile studies.

Finally, if, despite those efforts, we are still concerned about comprehension, we could require a test or tests—although designing such a test requires normative clarity about what kind of comprehension most matters.^{xvii}

These are proposals for ways of going forward in recruiting for risky cure studies while minimising the likelihood of non-autonomous enrolment. Would they be effective enough to quiet legitimate worries? That depends on how effective they are, which is ultimately a matter of empirical study. It also depends on the conceptual question of how exactly confident we have to be that a participant is enrolling fully autonomously. I believe these processes can be made adequately effective; nonetheless, it is worth repeating that as the risks increase, the required confidence does too.

RISK CEILINGS

As risks increase, the required confidence does too. This makes justifying riskier trials more difficult. Still, one might naturally wonder whether, in the present view, there is any *absolute* upper limit on permissible risk. Miller and Wertheimer, from whom we are drawing inspiration, suggest views like theirs are unlikely to support such a limit.¹ Why? Because people seem capable of autonomously choosing very serious risks. Miller and Wertheimer mention those who climb Everest, firefighters, and soldiers. We might add, more extremely, those who seek to free climb El Capitan or the heroes-in-the-moment who hug suicide bombers.^{15 16} In light of such examples, it may strike us that people can autonomously assume arbitrarily high risks, up to and including certain death. And this, in turn, can be a source of worry with respect to the view that respect for their autonomous choice exhausts the level of risk protection we owe participants.

Here is the first way that worry can be developed: by asking 'are we then committed to allowing absurdly risky trials to go forward *even when those trials are stupid and pointless (or otherwise lack social value)?*' After all, people can autonomously choose to do things that are risky, stupid and pointless—that is just how Alex Honnold's decision to free climb El Capitan strikes many. Suppose there was a trial with serious risk of death, whose only upshot was a treatment that might relieve mild headaches. Is the *only* restriction on the conduct of such a trial that we must be able to find a group of people idiosyncratic enough to genuinely value participation?^{xviii}

No. As trials get riskier, and as their objectives become more trivial, two things happen: it becomes harder and harder to find people for whom enrolment makes sense; and at the same time, the acceptable rate of type 1 errors goes down. In order to nevertheless maintain a tolerably low rate of type 1 error, we would need to install *substantial* procedural safeguards ensuring enrollees *really do* (quite surprisingly) value defeating

^{xvi}This example may be realistic. Cures will have to compete with lifelong ART in terms of cost-effectiveness and they may especially struggle on that front in low income contexts.³⁸

^{xvii}For related discussion, see Bromwich and Millum.³⁹

^{xviii}This example is taken from Jansen and Wall's critical discussion.⁴⁰

mild headaches enough to risk death. Those safeguards would not be free, as exercising social oversight requires the investment of social resources. For a trial that by stipulation has low-or-no social value, society may reasonably decline to invest those resources.

Thus, the view I have been developing is not committed to endorsing extreme, purposeless risks. Nonetheless, I do believe it is unlikely to supply an upper-risk limit in *all* cases, most notably, in those cases where the risks have significant social value. This is quite relevant to what could (or could not) be done in the name of developing an HIV cure.

Consider the ongoing Last Gift project. The Last Gift asks PLWHIV who have a terminal diagnosis (6 months or less) to agree to have samples taken during the remainder of their life, and, when they die, to have a rapid research autopsy performed. Due to the ethical and social complexity of end of life research, the Last Gift project has been conscientious in undertaking community consultation and centring patient participants in the research process. They also perform psychosocial research on participant experience; so far, the response has been overwhelmingly positive.¹⁷

The principle intervention in Last Gift occurs after death and is arguably *no* risk. But some researchers and ethicists have also endorsed the possibility of incorporating interventional research before death that would be decidedly more than minimal risk, including testing curative strategies.¹⁷ Carrying out that research at the end of life can help minimise the consequences of the attendant risks and toxicities, given that remaining time is anyway short. Indeed, although they do not say so, that would continue to be true even of research significantly more risky than the cure strategies they discuss.

Entertaining such possibilities may sound like an invitation to the abusive treatment of extremely vulnerable patients. But we should be careful to not superimpose our own values and concerns over theirs. One recent study of attitudes toward participating in cure research at the end of life, for instance, found that 39.6% of HIV-positive individuals surveyed were willing to shorten their lifespan between 0 and 4 weeks, and 31.2% were willing to shorten their lifespan by >4 weeks.¹⁸

We again cannot treat these survey results as interchangeable with actual enrolments. It may be easier to think warmly on making self-sacrifices for others when considered prospectively, from a long distance, but harder to do so when the appointed hour arrives. Social desirability bias may play a role. And further complicating things, in actual practice we rarely know exactly when the last 4 weeks of life have begun, let alone the last 6 months.

But still, consider more evidence, here in the form of quotes from an interview study of hospice patients who were rendered unable to participate in HIV research by their frail health¹⁹:

I feel like these last few weeks are wasted.
I wish I could do something else to help.
At least I could be doing something.

These quotes are offered as characteristic of all 12 interviews. If nothing else, they indicate that interest in participating in the end of life research need not evaporate as the end of life draws near (and may intensify). They also demonstrate a moving and eminently understandable search for meaning and desire to help.

It is also worth noting, as before, that *even if* the above evidence suggests a drastically inflated estimate of the number of people who could autonomously enrol in research that shortened or otherwise afflicted the end of their lives, that is still

compatible with there still being more than enough for certain kinds of high-value research to go forward. Depending on the study type, we may need only very few people to have the requisite motivations.

Now here it has been objected to me that in focusing on the end of life as a site for very risky research I have, in effect, cheated. Precisely because people at the end of life have less left to lose, the risks are not actually that high—making this not a genuine test case.

I have three responses. Two are essentially clerical. The first is that the standard party line is that research risks are *not* discounted according to length and quality of life remaining, so this objection assumes a heterodox accounting; nonetheless, I agree that the heterodox accounting is correct and so do not contest the point. The second is that even if the end of life cases are properly understood as not exceptionally high risk, they are still an area where high-value research could potentially go forward, but, I hazard, there are existing ethical concerns, so addressing those concerns seems independently worthwhile.

The third and more substantial response is this: okay, forget the end of life. Think of young healthy volunteers and consider a case which is genuinely high risk. ‘Challenge’ studies expose healthy volunteers to the disease for the purpose of gathering observations in a controlled setting. Suppose that someone proposed an HIV challenge study, under the rationale that doing so was scientifically necessary to learn crucial new information about the disease. Could *that* kind of study be permissible, if we found enough young people who were committed altruists?

I am willing to say that such a study could indeed be permissible. But the details are important. The research would need to be *valuable* enough that we could anticipate altruists actually entertaining it and enough that it actually made sense to invest the resources required to adequately oversee the process. Every caution already issued would still apply in full (I will not repeat my full recommendations from the last section). Meeting such cautions might be *very difficult*, so difficult that once everything is priced in the research no longer looks worthwhile. But then again, maybe not. It is hard to say in the abstract. In any case, I find these risks more worthwhile than those involved in climbing a mountain.

In the end, I am willing to defend the rejection of any *in principle* risk ceiling, over and above the one given by people’s values and the logistical difficulties of satisfactorily ascertaining them. Nonetheless, I close on a concessive note. The foregoing analysis has occurred under the assumption that such a risk ceiling is to be justified in terms of (il)legitimate paternalism: it is because we must protect people from their own potentially bad enrolment decisions. But, as I said in the introduction, this is one among many normative considerations potentially relevant to the conduct of research. My arguments are compatible with the thought that perhaps there is some other relevant consideration that must be given its due. I have already mentioned social trust and legal liability.¹²⁰ I cannot and anyway lack interest in enumerating every possible rationale. Rather, the point of the paper is just that if one wants to argue for a risk cap, with respect to HIV cures or elsewhere, here is one particular way *not* to do it: do not argue that we must protect human subjects from trials that genuinely advance their values.

Contributors RS is the sole contributor.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID ID

Robert Steel <http://orcid.org/0000-0003-4879-2070>

REFERENCES

- 1 Miller FG, Wertheimer A. Facing up to paternalism in research ethics. *Hastings Cent Rep* 2007;37(3):24–34.
- 2 Protection of human subjects, 45 C.F.R. §46.111, 2018. Available: <https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html> [Accessed cited 2019 April 29].
- 3 El-Sadr WM, Lundgren JD, Neaton JD, et al. Cd4+ count-guided interruption of antiretroviral treatment. *N Engl J Med* 2006;355(22):2283–96.
- 4 Julg B, Dee L, Ananworanich J, et al. Recommendations for analytical antiretroviral treatment interruptions in HIV research trials—report of a consensus meeting. *The Lancet HIV* 2019;6(4):e259–68.
- 5 Dubrocz G, Rakhmanina N. Antiretroviral therapy interruptions: impact on HIV treatment and transmission. *Hiv/aids* 2018;10:91–101.
- 6 Henderson GE, Peay HL, Kroon E, et al. Ethics of treatment interruption trials in HIV cure research: addressing the conundrum of risk/benefit assessment. *J Med Ethics* 2018;44(4):270–6.
- 7 Eyal N. What can the lived experience of participating in risky HIV cure-related studies establish? *J Med Ethics* 2018;44:277–8.
- 8 Namazi G, Fajnzylber JM, Aga E, et al. The control of HIV after antiretroviral medication pause (Champ) study: posttreatment controllers identified from 14 clinical studies. *J Infect Dis* 2018;218(12):1954–63.
- 9 Henderson GE, Waltz M, Meagher K, et al. Going off antiretroviral treatment in a closely monitored HIV “cure” trial: longitudinal assessments of acutely diagnosed trial participants and decliners. *J Int AIDS Soc* 2019;22(3):e25260.
- 10 Wendler D, Rid A. In defense of a social value requirement for clinical research. *Bioethics* 2017;31(2):77–86.
- 11 Kuritzkes DR. Why cure, why now? *J Med Ethics* 2017;43(2):67–70.
- 12 [ada.gov/HIV Home Page](http://ada.gov/HIV) [Internet]. Available: <https://www.ada.gov/hiv/> [Accessed 18 Apr 2019].
- 13 Simmons R, Kall M, Collins S, et al. A global survey of HIV-positive people’s attitudes towards cure research. *HIV Med* 2017;18(2):73–9.
- 14 Evans D. An activist’s argument that participant values should guide risk–benefit ratio calculations in HIV cure research. *J Med Ethics* 2017;43(2):100–3.
- 15 Magazine. How Alex Honnold made ‘the ultimate climb’—without a rope [Internet, 2019. Available: <https://www.nationalgeographic.com/magazine/2019/02/alex-honnold-made-ultimate-climb-el-capitan-without-rope/> [Accessed 12 Mar 2019].
- 16 Nordland R, Abed F. A Policeman’s Bear Hug Stops a Suicide Bomber From Killing More, 2017. Available: <https://www.nytimes.com/2017/11/16/world/asia/kabul-explosion-police.html> [Accessed 21 Feb 2019].
- 17 Dubé K, Gianella S, Concha-García S, et al. Ethical considerations for HIV cure-related research at the end of life. *BMC Med Ethics* 2018;19(1):83.
- 18 Prakash K, Gianella S, Dubé K, et al. Willingness to participate in HIV research at the end of life (EOL). *PLoS One* 2018;13(7):e0199670.
- 19 Gianella S, Taylor J, Brown TR, et al. Can research at the end of life be a useful tool to advance HIV cure? *AIDS* 2017;31(1):1–4.
- 20 Resnik DB. Limits on risks for healthy volunteers in biomedical research. *Theor Med Bioeth* 2012;33(2):137–49.
- 21 Eyal N. The benefit/risk ratio challenge in clinical research, and the case of HIV cure: an introduction. *J Med Ethics* 2017;43(2):65–6.
- 22 Miller PB, Weijer C. Rehabilitating equipoise. *Kennedy Inst Ethics J* 2003;13(2):93–118.
- 23 Weijer C, Miller PB. When are research risks reasonable in relation to anticipated benefits? *Nat Med* 2004;10(6):570–3.
- 24 London AJ. Reasonable risks in clinical research: a critique and a proposal for the integrative approach. *Stat Med* 2006;25(17):2869–85.
- 25 London AJ. Two dogmas of research ethics and the integrative approach to human-subjects research. *J Med Philos* 2007;32(2):99–116.
- 26 Anderson JA, Kimmelman J. Extending clinical equipoise to phase 1 trials involving patients: unresolved problems. *Kennedy Inst Ethics J* 2010;20(1):75–98.
- 27 Miller FG, Joffe S. Limits to research risks. *J Med Ethics* 2009;35(7):445–9.
- 28 Rid A, Wendler D. A framework for risk-benefit evaluations in biomedical research. *Kennedy Inst Ethics J* 2011;21(2):141–79.
- 29 Rajczi A. Making risk-benefit assessments of medical research protocols. *J Law Med Ethics* 2004;32(2):338–48.
- 30 Wen Y, Bar KJ, Li JZ. Lessons learned from HIV antiretroviral treatment interruption trials. *Curr Opin HIV AIDS* 2018;13(5):416–21.
- 31 Routy JP, Boulassel MR, Nicolette CA, et al. Assessing risk of a short-term antiretroviral therapy discontinuation as a read-out of viral control in Immune-based therapy. *J Med Virol* 2012;84(6):885–9.
- 32 Stunkel L, Grady C. More than the money: a review of the literature examining healthy volunteer motivations. *Contemp Clin Trials* 2011;32(3):342–52.
- 33 Grady C, Bedarida G, Sinaii N, et al. Motivations, enrollment decisions, and socio-demographic characteristics of healthy volunteers in phase 1 research. *Clin Trials* 2017;14(5):526–36.
- 34 Miller FG. Research ethics and misguided moral intuition. *J Law Med Ethics* 2004;32(1):111–6.
- 35 Emanuel EJ, Wendler D, Grady C. What makes clinical research ethical? *JAMA* 2000;283(20):2701–11.
- 36 Borducchi EN, Cabral C, Stephenson KE, et al. Ad26/MVA therapeutic vaccination with TLR7 stimulation in SIV-infected rhesus monkeys. *Nature* 2016;540(7632):284–7.
- 37 Largent E. For love and money: the need to rethink benefits in HIV cure studies. *J Med Ethics* 2017;43(2):96–9.
- 38 Freedberg KA, Sax PE. Improving on effective antiretroviral therapy: how good will a cure have to be? *J Med Ethics* 2017;43(2):71–3.
- 39 Bromwich D, Millum JR. Informed consent to HIV cure research. *J Med Ethics* 2017;43(2):108–13.
- 40 Jansen LA, Wall S. Reconsidering paternalism in clinical research. *Bioethics* 2018;32(1):50–8.