Gene Drive Mosquitoes: Ethical and Political Considerations

Athmeya Jayaram and Daniel Callies

Introduction

CRISPR "gene drives" are a promising and terrifying new technique in genetic modification. They allow scientists to modify the genes of organisms such that the modification is almost guaranteed to be passed on to the next generation, instead of the usual 50/50 chance. If an organism reproduces often enough, as mosquitoes do, then the modified organisms will make up a greater and greater share of the species until every wild type has been replaced with a modified organism. A gene drive can therefore change entire species with a single intervention.

The project to develop gene drives in mosquitoes is an effort to drive a particular modification through entire species of mosquitoes. The modification will make mosquitoes resistant to the malaria parasite, making them less likely to spread it to humans. If it works, this could solve the "last mile" problem of anti-malaria efforts. If we can't get treatments and bednets to certain areas, we can alter the mosquitoes there so that conventional methods aren't required. This could save millions of lives.

But, it could also worsen the health and environment of already vulnerable people. Genes that make mosquitoes resistant to malaria could also have other, unknown consequences for both mosquitoes and the humans they bite. Or, the malaria parasite might mutate to meet the challenge, presenting a possibly deadlier threat that renders conventional treatments ineffective. The risks of any new technology, especially a new genetic technology, are significant, and we can expect to see unintended consequences. The challenge for ethics, then, is to assess whether, given the risks, we should develop gene drive mosquitoes. And, if so, how we should design them, where we should release them, and who gets to make these decisions.

In what follows, we'll examine eight ethical concerns. Does mere research into gene drive mosquitoes place us on a slippery slope towards their use? Given the uncertainties involved in releasing gene drive constructs into the wild, might we be better off simply playing it safe and abandoning the gene drive project? In pursuing a gene drive approach to malaria eradication, are we ignoring the success we've achieved with conventional vector control methods in favor of a techno-fix? Does releasing gene drives into the wild fail to show nature the proper respect it deserves? Are we acting hubristically by attempting to gain ever more control over the natural environment? How should the public be involved in decisions surrounding research and development? How can we fairly distribute the benefits and burdens from gene drive mosquitoes? And is there a dual-use concern that gene drives could be used as weapons of destruction? These are the questions this chapter explores.

The Slippery Slope of Research

Gene drives are controversial. But it isn't just the actual use or release of gene drive constructs in the wild that causes concern. Some would argue that even research into the technology is something that should be abandoned. One of the most oft-cited worries about research is the slippery slope worry. The thought is that once research starts it will continue inexorably until gene drives are released outside of laboratories. That is, taking the first step of research will lead us inevitably down a slippery slope, at the bottom of which is the use of dangerous or morally fraught technology.

Slippery slope arguments all take the same form: If we allow *x*, then *y* will follow (the *empirical premise*). We don't want *y* to follow (the *normative premise*). Therefore, we shouldn't allow *x* (the *conclusion*). Being instances of *modus tolens*, these arguments are valid; their premises, if true, logically support their conclusions. The real question is whether these arguments are sound – whether, in addition to being valid, they, in fact, have true premises. And whether or not any particular slippery slope argument is sound will depend upon how the two premises are filled out and what relation they have to one another.

Consider an initial slippery slope worry about gene drive research:

> P1: If we research gene drive mosquitoes, this will lead to their release in the wild.
> P2: We don't want gene drive mosquitoes released in the wild.
> C3: Therefore, we shouldn't research gene drive mosquitoes.

As I mentioned above, this is a valid argument. But the soundness of the argument is in question. The first premise, the empirical premise, seems true. It seems very likely that research into gene drive mosquitoes will lead to their release in the wild. Maybe not tomorrow, maybe not this year (2021). But it seems likely that they will one day be released, and so the premise is likely true (though, we will have to wait to see about that). However, the truth-value of the second premise, the normative premise, is far less clear. That is, it isn't clear that we don't want to release gene drive mosquitos in the wild. It may very well be the case they turn out to be a highly effective tool with which to help eradicate malaria. Our assessment is that it is currently too early to tell whether the release of gene drive mosquitoes to eradicate malaria is something we should welcome or avoid. This is because, if successful, gene drive mosquitoes could confer immense public health benefits to some of the most disadvantaged populations of the globe. So, it is too early to say that we should avoid their release in the wild. For this first slippery slope argument, the first (empirical) premise is likely true, but the second (normative) premise is suspect, which makes the soundness of the argument as a whole suspect.

But consider a more dystopian slippery slope worry about gene drive research.

> P1: If we research gene drive mosquitoes, this will lead to a world full of gene-drive-modified crops, gene-drive-modified animals, and even gene-drive-modified humans.
> P2: We don't want a world full of gene-drive-modified crops, animals, and humans.
> C3: Therefore, we shouldn't research gene drive mosquitoes.

This more dystopian slippery slope argument is, like the first, a valid argument. If the premises are true, then the conclusion necessarily follows. There are some who would not be deterred by the second, normative premise in this argument. But equally true is that many would be. Let's grant for the moment the second premise is true – let us grant that a world in which gene drives are nearly omnipresent is something to be avoided. Notice that the first, empirical premise in this more dystopian argument is far more questionable. It is not at all clear that researching gene drive mosquitoes for malaria eradication will lead to a completely engineered world, with gene drive constructs being nearly omnipresent.

One of the hallmarks of a troubling slippery slope is when there is no clear point on the slope at which to draw a line. That is, a slippery slope worry is appropriate when the slope is, in fact, slippery, and there are no points on the slope to draw a principled distinction. But notice that there are multiple such points on the slope in question in this argument. There are clear lines between researching gene drive mosquitoes and using gene drives in humans. As it stands, this more dystopian slippery slope argument is also unsound; this time because the first, empirical premise is likely false.

What the slippery slope worry should push us towards is legitimate regulation of gene drives. With good regulation, we stand a better chance of safely and responsibly pursuing the benefits potentially associated with gene mosquitoes without worrying that this will lead to a world in which gene drives are omnipresent.

Precautionary Principle

Given that there is a non-zero chance that gene drives could cause significant environmental destruction, one might think the risk is simply not worthwhile. That is, one might think it better to err on the side of caution, and forgo the benefits that gene drives could deliver, in order to be certain that the destruction they could cause doesn't come about.

This kind of precautionary thinking is captured well by what has come to be known as the precautionary principle. The precautionary principle is a norm that aims to offer policy guidance in situations of uncertainty. The principle has been elaborated and endorsed by different governing bodies since at least the 1980s. The 1982 United Nations World Charter for Nature stated "Activities which are likely to pose a significant risk to nature shall be preceded by an exhaustive examination; their proponents shall demonstrate that expected benefits outweigh potential damage to nature, and where potential adverse effects are not fully understood, the activities should not proceed" (United Nations, 1982). Put slightly differently the 1992 United Nations Rio Declaration states "Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation" (United Nations, 1992b). And the world's leading treaty on biological diversity – the United Nations Convention on Biological Diversity – reads: "Where there is a threat of significant reduction or loss of biological diversity, lack of full scientific certainty should not be used as a reason for postponing measures to avoid or minimize such a threat" (United Nations, 1992a).

As should be clear from the preceding paragraph, there is no one canonical precautionary principle. Rather, the precautionary principle gets at the idea contained in the old adage of "better safe than sorry." Despite there being no singular version of the precautionary principle, Neil Manson has pointed out that each variant follows the same basic formula. According to Manson, there is always first a damage condition, second a knowledge condition, and finally a remedy. Take the formulation found in the Convention on Biological Diversity, for example. "Where there is a threat of significant reduction or loss of biological diversity [the damage condition], lack of full scientific certainty [the knowledge condition] should not be used as a reason for postponing measures to avoid or minimize such a threat [the remedy]" (United Nations, 1992a).

What does the precautionary principle, as elaborated in the Convention on Biological Diversity, imply for the gene drive mosquito malaria eradication project? Given that we can't be certain of the effects the gene drive mosquito project would have on ecosystems, we have to admit that there is a *threat* of a loss of biodiversity, even if the threat is a very low probability one. And the Convention on Biological Diversity notes that "lack of full scientific certainty should not be used as a reason for postponing measures to avoid or minimize such a threat." It would seem that the precautionary reasoning embedded in the Convention would preclude us from releasing gene drive mosquitos in the wild.

But many scholars, including Cass Sunstein (Sunstein, 2005), have noted how crippling – and, at times, incoherent – the precautionary principle is. The principle places an incredibly high burden of proof on new technologies. If we held all new technologies to the test that they must prove, beyond a reasonable doubt, that there is no threat of serious harm, then many things we currently rely upon— chlorine, airplanes, antibiotics, vaccines, X-rays, etc.—all would have been abandoned.

Ultimately, the advice of taking precaution, or attempting to avoid catastrophe, offers little to no guidance. The situation we are in is one in which the status quo is unacceptable. Millions of people are infected with malaria each year, and hundreds of thousands die from the disease. Simply recognizing that one ought to be cautious fails to deliver any sound policy recommendations. This being the case, it is unlikely that a crude version of the precautionary principle, like the wording found in the Convention on Biological Diversity, can help us reach a conclusion about the best course of action, all things considered.

Conventional Alternatives

Another ethical concern one might have about gene drives could come from a suspicion about technocratic approaches to social problems. Some worry that a gene drive approach to malaria might be a quintessential technological fix or "techno-fix." As Alvin Weinberg has pointed out, a techno-fix occurs when a technological or engineering approach is taken to solve a problem instead of pursuing a social or behavioral change (Weinberg, 1967). For instance, one could view gastric band – or lap band – surgery as a technological fix. Rather than an individual pursuing a demanding behavioral change, they could instead opt to address their weight and heart issues with a surgery that constricts the amount of food their stomach can take in, thus reducing their food intake. We might be worried about society pursuing these kinds of technological approaches to problems, rather than changing institutions or our collective behavior. With respect to the gene drive mosquito project, one could worry that we are opting for a technological solution rather than the more demanding avenue of investing fully in conventional malaria control methods and health care infrastructure in poorer parts of the world.

There are two responses to the techno-fix concern that are worth pointing out. First, as Christopher Preston has argued (albeit in a different context), the moral status of techno-fixes is not entirely clear (Preston, 2013). That is, while the word has a negative connotation (Rosner, 2004), it isn't clear that any particular "techno-fix" will be prima facie morally suspect. Whether or not we should have ethical reservations about a particular technological approach to a problem will depend upon the particulars of the technology, the magnitude of the problem it is attempting to address, and the specifics of the social and behavioral avenues that could also be pursued.

This leads directly into the second response to the techno-fix concern. It isn't entirely clear that there is a socio-behavioral approach to the problem of malaria that is being overlooked in favor of a gene drive approach. There are, of course, a host of conventional malaria control methods that have been rather successful at eliminating much of the disease burden. Bednets, indoor residual spraying, the elimination of standing water, and more widely dispersed access to health care have certainly reduced the burden of malaria in many parts of the world (Callies, 2020). For example, malaria has been eradicated from countries like Algeria to Singapore with conventional vector control methods and investment in health care infrastructure. (World Health Organization 2017).

But it is also important to note that countries like Algeria and Singapore did not face anywhere near the disease burden faced in sub-Saharan Africa. The WHO (2019: 2) writes, "…even with our most optimistic scenarios and projections, we face an unavoidable fact: using current tools, we will still have 11 million cases of malaria in Africa in 2050." So, it isn't entirely clear that there is an easy social change that could bring about the same potential reduction in malarial burden that could be achieved through gene drives.

One should also recognize that there are important differences between eliminating malaria with gene drives and doing so with increased investment in conventional methods. On the one hand, investing in health care infrastructure throughout sub-Saharan Africa would certainly help eliminate the burden of malaria. And this would have significant corollary benefits. But this would also be massively costly, and funding for malaria control and general development in this part of the world is already in short supply. The WHO estimated that 4.4 billion dollars was needed for conventional malaria control efforts in 2017, and yet only 3.1 billion dollars was provided. And, of course, this WHO estimate does not include massive investments in healthcare infrastructure.

Furthermore, some conventional control methods – like the draining of wetlands and the wide use of insecticides – have significantly deleterious ecological effects. Draining wetlands to prevent mosquitoes from breeding destroys entire ecosystems, and insecticides are similarly indiscriminate with the kinds of invertebrates they kill. So, more investment in conventional malaria control methods and much greater investment in health care infrastructure would certainly be desirable. But it may not be a sufficient response to the problem. And if there is not a clear socio-behavioral avenue to eliminating malaria, it's not clear how much weight the techno-fix worry should be given.

Environmental Ethics

Environmental ethics often asks us to think about things from outside of our own perspective. Imagine we were to go ahead with the gene drive mosquito project and everything were to go exactly as planned. Regardless of whether we used a suppression drive or a replacement drive, imagine we were able to reduce malaria to zero and there were minimal adverse side-effects. Some environmental philosophers would still claim that we could be doing something wrong or failing to show nature the proper respect it deserves. There are different ethical outlooks in environmental ethics – with three being the most common.

Perhaps the most common is the anthropocentric outlook. Anthropocentrism, within environmental ethics, is the view that value is human-centered. Humans are the kinds of things that have intrinsic or final value, and anything else with value has its value as a means to some human end. But there are other outlooks on value as well. The biocentric outlook would place intrinsic or final value not just in humans, but in all individual forms of life. Any individual organism that counts as living is something that has intrinsic or final value. Someone like Holmes Rolston III would note that each species is a unique embodiment of the evolutionary process, and that this uniqueness confers intrinsic value. The ecocentric outlook would place intrinsic value not just in humans, and not just in individual forms of life, but in collectives of life or ecosystems. Ultimately, the kinds of things that have intrinsic or final value are natural communities or ecosystems. Ecocentrists, like Aldo Leopold, might note that protecting the value found in individuals could lead to significant harm to groups or ecosystems. For instance, individual members of invasive species can often damage ecosystem integrity. A biocentrist, someone who places value in individual organisms, may object to sacrificing one individual organism even for the sake of the whole. Whereas an ecocentrist, like Leopold, would focus on the group and say eliminating the one for the sake of the group is permissible, and sometimes even obligatory.

Each of these environmental outlooks has something to say for them. And it is not our intention to rehearse the merits and demerits of each here. Rather we want to note that the gene drive mosquito project could be seen as troubling from any of the three outlooks. It is easy to see how biocentrists and ecocentrists could take issue with the malaria eradication project were to it travel down the population suppression route. The biocentrist would object to the eradication of the particular individual specimens and the ecocentrist would worry about the removal of the species from the ecosystem. But even the most restrictive of these ethics, anthropocentrism, could find the gene drive malaria project troubling. Anthropocentrism identifies great value in nature, and that is in the form of nature's instrumental value. That is, an anthropocentric environmental ethic will still see nature as valuable because it is valuable *to us*.

There are at least two different aspects of nature that we value. On the one hand, we value nature's *autonomy*. The eighteenth-century philosopher John Stuart Mill noted that "nature" often refers to that which "takes place without the agency, or without the voluntary and intentional agency, of man"(Mill, 2008). The autonomy or independence of nature is something we value. In addition to its autonomy, there is a second sense of nature and naturalness that matters to us. For example, according to a founding document of the US National Park System, the primary goal of management should be that the “biotic associations within each park be maintained…. A national park should represent a vignette of primitive America” (Leopold et al. 1963). The key takeaway from the Leopold Report is that a sense of nature or naturalness is had in the *composition* of an ecosystem, of it bearing historical resemblance to the past. Following Gregory Cole and David Aplet, we can refer to this as nature’s *ecological condition* (Cole & Aplet, 2010).

Importantly, these aren’t binary aspects of nature. Rather, they are continua. So, we can imagine the continuum of autonomy ranging from a completely human controlled environment to a completely autonomous or self-willed environment. And we can imagine the continuum of ecological condition ranging from a completely novel ecosystem that bears little to no fidelity to the past, to a completely pristine ecosystem that is exactly as it would have been at some point in the past. We tend to value nature the more autonomous or wild it is, and we similarly tend to value nature the more its ecological condition maintains fidelity to some point in our history.

If we were to use a suppression drive to drive down or eliminate a particular mosquito population, we would be changing nature’s ecological condition; we would be changing the composition of certain ecosystems, and doing so intentionally, thus diminishing nature’s naturalness. Similarly, if we were to use a replacement drive, we would be affecting nature’s autonomy; we would be intervening and controlling the organisms within an ecosystem, even if they maintained a kind of historical fidelity to previous ecosystems. Thus, even from an anthropocentric outlook, either malaria eradication approach will result in a loss of natural value of some kind.

Whether one subscribes to an anthropocentric, biocentric, or ecocentric outlook, either gene drive approach would carry with it a loss of some kind of value. But, importantly, it is doubtful that an anthropocentrist, biocentrist, or ecocentrist would think that this loss of value would *necessarily* make the agenda impermissible. Anthropocentrists find value in humans, and insofar as the gene drive malaria eradication project would produce significant anthropogenic benefits, the loss in natural value could be easily justified. And even biocentrists and ecocentrists note that our duties to the natural world are not absolute. Perhaps the most famous biocentrist of all time, Holmes Rolston III wrote: “The duty to species can be overridden, for example with pests or disease organisms.” And when asked about mosquitoes in particular, the famous ecologist E. O. Wilson said: “I would gladly throw the switch and be the executioner myself.”

Basically, even if individual organisms or ecosystems have intrinsic value, so, too, do humans. Furthermore, organisms and ecosystems may have non-deontic intrinsic value (by which I mean value that doesn’t create duties). And given the death and suffering engendered by malaria, the loss of natural value may be insufficient to make the agenda impermissible.

Hubris

The control that we can exhibit over the natural environment seems to be accelerating exponentially. And the ability to push specific alleles through a population and then subsequent generations is awesome, in the literal sense of the word. But one might wonder whether we are not overestimating our ability to control nature. Many ethical theories tell us to think about whether our actions are good or bad, permissible or impermissible. But a tradition dating back to Hellenic Greece and Aristotle pushes us to focus instead on the kinds of character traits our actions exhibit. Known as *virtue ethics* (Annas, 2005; Hursthouse, 2001; Nussbaum, 2013), this ethical theory places focus not so much on what we should do, but on how we ought to be (Driver, 2017).

In focusing on how we ought to be, virtue ethics pushes us to develop positive character traits – virtues – and avoid negative character traits – vices. There are myriad virtues and vices worthy of consideration, but one has been of particular focus when it comes to environmental matters: hubris. Hubris describes a dangerous overconfidence in one's abilities. One need not look too far to find examples in which a hubristic attitude has led to environmental disaster.

Many households throughout the U.S. used DDT as a pesticide by the middle of the 20th century, and the U.S. military relied upon the chemical compound as a defoliant during the Vietnam War. It wasn't until the 1960's that the public became aware of DDT's harmful effects. The publication of Rachel Carson's *Silent Spring* brought to the fore the negative side-effect DDT had on the environment, specifically the eggshell-thinning effect it had on many North American birds, including the iconic bald eagle (Carson, 2002). One might be tempted to refer to the unforeseen negative side-effects of DDT and other environmental interventions not has manifestations of hubris, but rather as instances of insufficient risk-assessment prior to the use of novel technologies. But Hoffmann-Riem and Wynne note that "unanticipated effects of novel technologies are not just possible but *probable*" (Hoffmann-Riem & Wynne, 2002). What this means with respect to the gene drive malaria eradication project is that we should count on there being unanticipated side-effects. Though, it seems clear that we cannot have a good grasp of their magnitude *ex ante*.

But while we can count on there being unanticipated adverse side effects from gene drives, it isn't clear whether those adverse side-effects will be as detrimental as the malaria they try to prevent or the conventional control methods already in use, like the use of undifferentiating insecticides. While we'll want to adopt a humble attitude and avoid hubris when it comes to our attempts to control nature, there are other character traits equally important in the virtue ethics tradition. Adopting and acting in accordance with the virtues of compassion and beneficence, for example, should push us to refuse to accept the status quo in which hundreds of thousands of undeserving individuals fall victim to malaria each year. It is clear that we will want to avoid hubris and exemplify compassion and beneficence. But, and this is a common objection to virtue theory, it isn't entirely clear what virtue theory recommends as the right course of action, recognition of these virtues and vices notwithstanding.

Public Participation

Amid all the questions over how to weigh and distribute the risks and benefits of gene drive mosquitoes (GDMs), perhaps the most important question is: who gets to make these decisions? Who gets to decide whether we develop or release GDMs, where we release them, or how they should be designed?

Some would argue that the decisions should be made by experts, based on the expected value of the outcomes or on the demands of justice. However, there are two reasons to think that public participation should carry significant, or even decisive, weight in the final choice. First, experts may not have much of an advantage in making these decisions. The decision to release GDMs involves weighing unknown risks and potentially incommensurable values, so the public may be a better, and more legitimate, source of answers to such questions.

And second, the nature of the risks and benefits may trigger an ethical requirement for public influence (Callies, 2019). That is the claim of the "all-affected principle," which states that the people who are relevantly affected by a political decision should have influence over that decision. For example, imagine a country's legislature decides that the safest place to locate a nuclear power plant is near its border. Perhaps the legislature even fairly considers the interests of people across the border, whatever that entails. Still, it seems wrong that people on only one side of the border get to influence the decision, when both sides are equally affected.

The all-affected principle has an intuitive appeal, but it leaves open two crucial questions: (1) What does it mean to be "relevantly affected"? And, (2) How much influence does it entitle you to?

To answer the first question, we can consider what kind of effect would entitle me to participation rights. Clearly, there are some kinds of effects that would *not* entitle me to any influence. For instance, if Germany decided to require a visitors' visa for Americans, it may annoy me, but the effect is too trivial for Germany to be required to get my input. On the other hand, if I were a refugee seeking asylum in Germany, its decision would profoundly affect my life, so the all-affected principle would grant refugees some influence over that decision.

One theory of what it means for a decision to "relevantly affect" you is, therefore, that your basic interests are at stake – interests like security, nutrition, health, and education (Song, 2012). This is not the only plausible conception of "relevant effect." One might argue that decisions that affect the environment (like releasing GDMs) do not affect my basic interests, but still trigger a right to participate. For example, the decision to preserve or eliminate an animal species does not affect my basic interests, but it may still concern my value for nature in a way that gives me a right to weigh in. Nevertheless, we will focus here on the "basic interests" account as perhaps the strongest case for generating a right to participate.

People's basic interests are certainly at stake in the decision to develop and release GDMs. A best-case release of GDMs will remove the threat of malaria and improve health outcomes. A worst-case release of GDMs may worsen health outcomes or harm the environment in ways that affect lifestyles and livelihoods.

However, whose interests are at stake, and how much, depends on further complications in the theory, such as: should we consider any *possible* outcome of the decision to release GDMs or should we weight the outcomes according to probability (Arrhenius, 2018)? Considering any possible outcome means that people who live in *any* mosquito habitat may be affected by releasing GDMs. If so, we might have to include half the globe in the decision-making process.

Weighing the outcomes according to probability, on the other hand, results in tiers of influence. The stakes would remain roughly the same everywhere, since the same basic interests are affected. But, we would multiply the stakes by the probability that they are affected, which would be greatest closest to the initial release site and diminish as you get farther away. At some distance away, the probability may fall so low that we can consider the effect to be trivial, which would not trigger any participation requirement.

There are three major caveats to this analysis, however. First, if there is a slippery slope from research to initial deployment to widespread imposition, then the decision to research GDMs would significantly affect the decision to release them widely. Second, even if different communities are given an independent chance to decide on GDMs for themselves, research decisions affect the *kinds* of GDMs that they can decide upon. So, the malaria-suffering world may be entitled to a say in how GDMs are designed. And third, developing GDMs will make everyone who lives in mosquito habitats more susceptible to the malicious use of GDMs, which we discuss in the “dual use” section. All these factors may entitle the mosquito-suffering world to an influence over research decisions, as well as deployment.

There are other ways of fleshing out the all-affected theory but, however you do it, participation rights are unlikely to be *exclusive* to either the residents near the initial release site or to the citizens of a single country. The theory is also likely to grant more influence to people in malaria-suffering or mosquito-ridden areas, perhaps even over research decisions. Finally, if we accept the “basic interests” account of who is affected, then we should grant extra influence to the poor, whose life-outcomes will vary most based on the success or failure of GDMs.

If those are the groups who ought to influence the decision, the next question is: how much influence should they have? There are two main views. The first is that everyone who has a stake above a certain threshold should have an equal influence on the decision, which is presumably an equal vote (Goodin, 2007). It is difficult to justify this option if we measure stakes in relation to the probability of outcomes, since different groups would then have very different stakes but the same amount of influence.

Instead, we may opt for Brighouse and Fleurbaey’s view that people should have influence in proportion to their stake in the decision (Brighouse & Fleurbaey, 2010). One way to institutionalize that is by giving everyone a vote that is weighted by their stake in the decision, but that is not the only way. Many defenders of the all-affected principle are careful to note that the right to influence does not mean the right to a vote. We can offer influence in many other forms – consultation rights, the right to participate in deliberation, or even the right to relocate – depending on what would best protect people’s interests or motivate their participation. On this view, the all-affected principle supplies the ideal, which we realize through institutional design and community engagement.

Distributive Justice

The primary purpose of GDMs is to benefit the developing world, so it may seem odd to worry about whether such a project is fair for the intended beneficiaries. However, because GDMs come with significant risks, and those risks fall disproportionately on some groups, we must consider whether they are distributed fairly.

One way to justify assigning a disproportionate risk to some is by citing the beneficial consequences for the whole. For example, if we initially release GDMs on an island, then we could lower the total risk of mortality to humanity, so the disproportionate risk to those islanders might be justified. Or, for that matter, if we release them in India, we could reduce the risk of mortality to South Americans. However, non-consequentialists would argue that the overall benefits do not justify assigning some people more than their fair share of risk.

Instead, we might argue that the benefits *to the* islanders (or the Indians) are sufficient to justify the risks. After all, if GDMs work, they could eradicate malaria, saving many people from death and the high costs of illness. On one hand, this seems like an important contribution to fairness; it is certainly worse to ask people to take on disproportionate risk if they see no benefit from it. But, even on an optimistic view of the risks, the benefits do not seem sufficient to justify the risks here. Consider that, if GDMs work as hoped, then they could soon be used everywhere, which means that everyone would soon enjoy the same benefit. There may be some value to getting that benefit sooner, but it is unlikely to justify bearing the brunt of the risk.

If the benefits alone do not justify the risk, then we might seek the consent of those who bear the risks. For instance, if the islanders vote to host the GDMs, then we have reason to think that they see the benefits as outweighing the costs and are, therefore, not unfairly burdened. However, consent loses its justificatory power when the conditions under which it is given are unfair. To the extent that the health or socioeconomic condition of the islanders is a product of previous injustice, the consent is not freely given and doesn't justify the disproportionate risk. Consent may still be a necessary condition, but a concern about background inequality makes it all the more important that the risks from GDMs are distributed fairly.

So, let us consider two theories of how to fairly distribute the risks: egalitarianism and prioritarianism. An egalitarian theory would require that the risks be distributed equally in some sense. In the case of GDMs, however, this is unappealing because it would greatly increase the total risk to humanity. For example, equality may mean that we release GDMs everywhere at once. Or, it may mean that we choose the initial release site by a lottery where each citizen has an equal chance of selection, making the most populous areas the most likely release sites.

A prioritarian theory, on the other hand, allows for unequal distributions as long as they are to the benefit of the least advantaged group. In this case, either the country (e.g. India) or the community around the release site would probably count as least advantaged among its peers. What arrangement would be to the advantage of the country or community that bears the risk?

One answer is to release GDMs in a developed country or a wealthy community. It would certainly be to the benefit of the least advantaged to find out if GDMs worked, without bearing any of the risk. Testing on more advantaged groups also alleviates our worries about consent and exploitation; the consent of rich communities will not be a product of background injustice or desperation. The problem with this strategy is that developed countries may not be scientifically suitable as test sites, and rich islanders are likely to block any such test. Neither development would be to the benefit of the least advantaged.

Instead, we could pose a similar question hypothetically: what it would it take for the rich to bear the risks from GDMs? This is essentially the strategy behind John Rawls' thought experiment, "the original position." In the original position, we consider the distribution of benefits and burdens in a society *as if* we did not know which social group we belonged to. If a reasonable person would be comfortable with belonging to any social group in a distributive scheme, then that distributive scheme is justified to all social groups.

So, the question becomes: what would it take for us to be indifferent to whether or not we belonged to the risk-bearing group? One possible answer is that nothing could make us accept the risks of GDMs. If the risks are high enough, a reasonable person might reject any arrangement under which she could have to bear those risks.

On the other hand, there might be conditions under which we *would* be indifferent between the risky and non-risky positions. For one, we would probably insist that every reasonable precaution was taken to minimize the risks of GDMs. And we might also require compensation for bearing those risks. The compensation might be a form of insurance for any harmful consequences, such as guaranteed medical care or relocation expenses. Or, it might be compensation for bearing the risks, whether or not the harms come to pass.

Of course, compensation is a controversial solution to this kind of injustice. Those who favor it argue that it gives all decision-makers an incentive to minimize risks. Without a compensation requirement, many groups stand to benefit from releasing GDMs without bearing any risk, such as the malaria-suffering nations far from the release site, or the developers and patent-owners of gene drive technology. But, if the beneficiaries have to compensate the risk-bearers, then the former have a reason to minimize the risks.

However, those who oppose compensation raise a series of difficult moral and practical questions: Is it fair for the rich to pay others to risk their health? Does it further undermine community consent if the community consents in order to receive compensation? What is the right level of compensation for risks that are hard to estimate or even anticipate? How can compensation address risks to future generations? Still, if GDMs are worth the risk, and the risk must be distributed unequally, then compensation, consent and precautions are our only tools to justify that inequality (Hayenhjelm, 2012).

The Dual Use Dilemma

The "dual use dilemma" is an ethical problem for scientists, administrators, grantmakers and all others involved in supporting research that could be used for great good or serious harm. The typical historical example is the research into nuclear fission, which scientists knew could be used to both generate energy and to make bombs. The dilemma is generally discussed today in connection with research into novel pathogens, which help prepare us for future diseases but also make their accidental or malicious spread more likely. Research into GDMs presents a similar problem. GDMs can be designed to *prevent* disease by making mosquitoes more resistant to spreading malaria, or they can be designed to *spread* disease by making them better hosts for malaria (or something more deadly).

Of course, most human creations can be used for both good and ill, so it is important to define the dual use dilemma more precisely. Thomas Douglas offers the following definition, which I have slightly amended (Douglas, 2013). A dual use dilemma arises when: (1) an agent is deciding whether to contribute to a scientific output; (2) That output could be used to produce *significant* benefit or harm; (3) The intended and primary use of the output is for good; (4) Increasing the likelihood of the good use also increases the likelihood of the bad use; (5) There's no option to create an output that's useful but harmless; and (6) The potential benefit and harm are of comparable moral worth.

Given that GDMs fit the conditions for a dual use dilemma, is it ethical to pursue them? A natural way to answer that question is to compare the expected benefits of GDMs with the expected harms; are GDMs likely to be a net benefit to the world? The problem is that it's very hard to estimate GDMs' expected benefits or harms before releasing them. We don't know whether and how successful they will be in reducing the spread of disease, what kinds of known and unknown harms might result, and how likely they are to be used for malicious purposes.

One important difference between GDMs and dual use projects like pathogens is that, with GDMs, we at least know that there's a problem to be solved; millions of people die from malaria and dengue every year. This stands in contrast to pathogen research, where it is hard to know whether there will even be a problem to solve. What are the chances that, for instance, a more transmissible airborne smallpox will ever develop, accidentally or maliciously?

Still, it is hard to estimate the overall expected value of GDMs, which requires an understanding of infectious disease, local environments, decision theory, and security risk, in addition to transgenic mosquitoes. So, however we estimate the expected value, we can at least conclude that it will require a variety of experts working together, perhaps, as some have proposed, in a biosecurity review board.

Rather than attempt any rough calculations of expected value here, we will briefly explore the ethical consequences of either a positive or negative result. Assume, first, that we arrive at a positive expected value; that is, we expect GDMs to do more good than harm. That is a strong *pro tanto* reason to develop them. But, can developing them still be wrong?

One could argue that, because their developers can foresee GDMs being used for harm, it is wrong to facilitate that harm, even if GDMs could also prevent a greater number of deaths. The typical response to this is to draw a distinction between foreseeing and intending a result. According to this response, it is permissible to develop GDMs if one merely foresees that they will be used for harm, as long as one does not intend that harm. For example, it is intuitively permissible for the military to bomb a weapons factory, even if it foresees that some civilians will die as a result. It is impermissible, however, for a terrorist to *target* the same number of civilians because, in that case, the terrorist intends their deaths.

Even if one accepts this distinction, however, it is important to note a relevant difference between the bombing case and GDMs. In the bombing case, there is (we assume) no way to mitigate the foreseeable casualties; once the bomb is released, the destruction follows. However, there are a number of options to possibly limit the harmful potential of GDMs, such as by developing reversible or self-limiting drives. If gene drive developers were to fail to mitigate foreseeable harm when possible, then they may be morally responsible for that harm, even if they did not intend it (Uniacke, 2013). And this could be true even if GDMs, overall, do more good than harm.

Now, let us consider the alternative. If we expected GDMs to do more harm than good; can it still be permissible to create them? Here is one argument that might absolve its creators of responsibility. They might say: “If I didn’t do it, someone else would have.” Since we can assume that *someone* would have created GDMs regardless, their actual creators are not morally responsible for their harmful effects.

However, this defense seems mistaken from either of the two major ethical perspectives. From a consequentialist perspective, we judge ethical actions by their consequences, and we can grant that GDMs would exist no matter what their creators did. Still, the creators’ actions *do* have consequences. If the creators had worked slowly or poorly, it would have at least delayed harmful consequences of GDMs.

And, from a nonconsequentialist perspective, it *does* matter if you contribute to a harmful result, even if the result would have been the same either way. Jonathan Glover names this principle after Alexandr Solzhenitsyn, who wrote: “Let the lie come into the world, even dominate the world, but not through me” (Glover & Scott-Taggart, 1975).

Conclusion

As should be clear from this discussion, making ethical choices on emerging technologies with profound social and environmental effects is both challenging and essential. The aim of this chapter has been to introduce some of the ethical concerns raised by the gene drive mosquito project. We’ve indicated where we think the balance of reason lies with respect to some of these issues, but most of them will require a closer examination of the arguments, and a broader discussion of our values. We hope this chapter will encourage those discussions among ethicists, scientists, and the public.

References

Annas, J. (2005). Virtue Ethics. In D. Copp (Ed.), *The Oxford Handbook of Ethical Theory* (pp. 515–536). Oxford University Press. https://doi.org/10.1093/0195147790.003.0019

Arrhenius, G. (2018). The Democratic Boundary Problem Reconstructed. *Ethics, Politics & Society*, *1*, 34. https://doi.org/10.21814/eps.1.1.52

Brighouse, H., & Fleurbaey, M. (2010). Democracy and Proportionality. *Journal of Political Philosophy*, *18*(2), 137–155. https://doi.org/10.1111/j.1467-9760.2008.00316.x

Callies, D. E. (2019). The ethical landscape of gene drive research. *Bioethics*. https://doi.org/10.1111/bioe.12640

Callies, D. E. (2020). Bednets or Biotechnology: To Rescue Current Persons or Research for the Future? *Fudan Journal of the Humanities and Social Sciences*, *13*(4), 559–572. https://doi.org/10.1007/s40647-020-00290-7

Carson, R. (2002). *Silent spring* (40th anniversary ed., 1st Mariner Books ed). Houghton Mifflin.

Cole, D. N., & Aplet, G. H. (2010). The Trouble with Naturalness: Rethinking Park and Wilderness Goals. In D. N. Cole & L. Yung (Eds.), *Beyond Naturalness: Rethinking Park and Wilderness Stewardship in an Era of Rapid Change.* (pp. 12–29). Island Press. http://qut.eblib.com.au/patron/FullRecord.aspx?p=3317526

Douglas, T. (2013). An Expected-Value Approach to the Dual-Use Problem. In B. Rappert & M. J. Selgelid (Eds.), *On the Dual Uses of Science and Ethics: Principles, Practices, and Prospects*. ANU Press. https://doi.org/10.22459/DUSE.12.2013.09

Driver, J. (2017). Virtue Ethics. In S. M. Cahn (Ed.), *Exploring ethics: An introductory anthology* (Fourth Edition, pp. 145–147). Oxford University Press.

Glover, J., & Scott-Taggart, M. (1975). "It Makes No Difference Whether or Not I Do It." *Aristotelian Society Supplementary Volume*, *49*(1), 171–210. https://doi.org/10.1093/aristoteliansupp/49.1.171

Goodin, R. E. (2007). Enfranchising all affected interests, and its alternatives. *Philosophy & Public Affairs*, *35*(1), 40–68.

Hayenhjelm, M. (2012). What Is a Fair Distribution of Risk? In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of Risk Theory* (pp. 909–929). Springer Netherlands. https://doi.org/10.1007/978-94-007-1433-5_36

Hoffmann-Riem, H., & Wynne, B. (2002). In Risk Assessment, One Has to Admit Ignorance. *Nature*, *416*(6877), 123.

Hursthouse, R. (2001). *On virtue ethics*. Oxford University Press.

Mill, J. S. (2008). *On liberty and other essays*. Oxford University Press.

Nussbaum, M. C. (2013). Non-Relative Virtues: An Aristotelian Approach. In R. Shafer-Landau (Ed.), *Ethical theory: An anthology* (2nd ed, pp. 630–644). Wiley-Blackwell.

Preston, C. J. (2013). Ethics and geoengineering: Reviewing the moral issues raised by solar radiation management and carbon dioxide removal: Ethics & geoengineering. *Wiley Interdisciplinary Reviews: Climate Change*, *4*(1), 23–37. https://doi.org/10.1002/wcc.198

Rawls, J. (1999). *A Theory of Justice*. Harvard University Press.

Rosner, L. (Ed.). (2004). *The technological fix: How people use technology to create and solve problems*. Routledge.

Song, S. (2012). The boundary problem in democratic theory: Why the demos should be bounded by the state. *International Theory*, *4*(1), 39–68. https://doi.org/10.1017/S1752971911000248

Sunstein, C. R. (2005). *Laws of Fear: Beyond the Precautionary Principle*. Cambridge University Press.

Uniacke, S. (2013). The Doctrine of Double Effect and the Ethics of Dual Use. In B. Rappert & M. J. Selgelid (Eds.), *On the Dual Uses of Science and Ethics: Principles, Practices, and Prospects*. ANU Press. https://doi.org/10.22459/DUSE.12.2013.10

United Nations. (1982). *World Charter for Nature*. http://www.un.org/documents/ga/res/37/a37r007.htm

United Nations. (1992a). *Convention on Biological Diversity* (p. Preamble). http://unfccc.int/essential_background/convention/items/6036.php

United Nations. (1992b). *Rio Declaration on the Environment and Development* (p. Principle 15). http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm

Weinberg, A. (1967). *Reflections on Big Science*. The MIT Press.