We Are Not Pontius Pilate: Acknowledging Ethics and Policy

James Alexander Hughes
St. Francis Xavier University
Computer Science
Antigonish, NS, Canada
jhughes@stfx.ca

William Hannah
St. Francis Xavier University
Antigonish, NS, Canada
bhannah@stfx.ca

Peter Kikkert
St. Francis Xavier University
Brian Mulroney Institute of Government
Antigonish, NS, Canada
pkikkert@stfx.ca

Barry MacKenzie
St. Francis Xavier University
History
Antigonish, NS, Canada
bmackenz@stfx.ca

Wendy Ashlock
Ashlock & McGuinness Cons. Inc.
Guelph, ON, Canada
washlock@ashlockandmcguinness.com

Sheridan Houghten
Brock University
Computer Science
St. Catharines, ON, Canada
shoughten@brocku.ca

Daniel Ashlock
University of Guelph
Mathematics and Statistics
Guelph, ON, Canada
dashlock@uoguelph.ca

Matthew Stoodley
University of Guelph
Bioinformatics
Guelph, ON, Canada
mstoodle@uoguelph.ca

Michael Dubé
Brock University
Computer Science
St. Catharines, ON, Canada
md12ol@brocku.ca

Rachel Brown
University of Guelph
Bioinformatics
Guelph, ON, Canada
brownr@uoguelph.ca

Amanda Saunders
University of Guelph
Bioinformatics
Guelph, ON, Canada
asaunder@uoguelph.ca

***Abstract*—A new AI system is being developed to optimize vaccination strategies based on the structure and shape of a community's social contact network. The technology is minimally constrained and not bound by preconceived notions or human biases. With this come novel *outside the box* strategies; however, the system is only capable of optimizing what it is instructed to optimize, and does not consider any ethical or political concerns.**

With the growing concern for systematic discrimination as a result of artificial intelligence, we acknowledge a number of relevant issues that may arise as a consequence of our new technology and categorize them into three classes. We also introduce four normative ethical approaches that are used as a framework for decision-making.

Despite the focus on vaccination strategies, our goal is to improve the discussions surrounding public concern and trust over artificial intelligence and demonstrate that artificial intelligence practitioners are addressing these concerns.

***Index Terms*—Artificial Intelligence; COVID-19; Epidemic; Ethics; Optimization; Pandemic; Policy; SARS-CoV-2; Simulation; Vaccinations.**

I. Introduction

The World Health Organization, the Centers for Disease Control and Prevention (CDC), and various other national and international organizations have existing procedures and guidelines for determining how and who to prioritize when vaccinating a population [2], [3]. These guidelines include extensive risk/benefit analysis, health economics, implementation issues, population values, ethics, demographic risks, and an emphasis on at-risk populations such as front-line workers. In addition to their intuitive value, these guidelines are extremely important and have been demonstrated to be effective in practice.

We are developing a system to provide additional details and guidelines based on the structure, shape, topology and properties of the population, how the population is distributed physically, and the current state of the pandemic disease. Artificial intelligence (AI) is used to automate the discovery of rules and vaccination strategies based on the structure of a given population. The goal is to provide public health officials and decision makers with additional information to augment existing practices, to ultimately improve the quality of response and reduce the overall impact of a given disease.

The system being developed is designed to be applicable to any disease; however, given the situation with the virus SARS-CoV-2, the work has gained attention from media and government officials [12], [13], [23]. During these conversations, many critically important ethical and political questions related to our work have arisen. These questions were not considered or even anticipated during the initial system design, but they are undeniably paramount with respect to the implementation and effectiveness of our scientific contribution. These questions were neither intentionally missed nor ignored out of malice. Rather, they were missed as they do not fall within our expertise, a consequence of our lack of experience in the relevant subject matter. Regardless of the intent, the system is being created and will have real world implications.

Whether such questions are addressed, missed, or willfully ignored, AI practitioners may be accused of "passing the buck" and contributing to and perpetuating systematic discrimination

978-1-7281-2547-3/20/$31.00

and biases[1]. This, unfortunately, only adds to public mistrust of AI in the realm of real world problems having a direct impact on many lives. Public trust in this technology is critical [24], especially within the context of this work and an already politicized virus — refer to the *Twitter* thread related to this work to see typical examples of public mistrust in AI and of the politically charged atmosphere concerning SARS-CoV-2 vaccines [51]. According to the *Angus Reid Institute*, as of August 4, 2020, only 46% of Canadians would get a vaccine as soon as it became available, with an additional 32% planning on waiting [26]. With public mistrust of AI, the proportion of the population not willing to receive a vaccine may increase.

The intentional design of our system's search has an additional complication when it comes to the ethics and policy around prioritizing individuals for vaccines. One of the major *benefits* of using AI for discovering vaccination strategies is that it lacks preconceived notions; it can *think outside the box* while performing the search. Although this enables the discovery of interesting and novel strategies, it also means that when the system synthesizes the strategies, it is *not bound by human biases*. While generally seen as a strength, this also means that human ethics and social or cultural norms are not incorporated. Do note however, bias may be incorporated into the systems while designing the system, selecting the data, and interpreting result.

The distribution of vaccines is a complex ethical and policy issue, as a strong case can be made to prioritize many different groups. For example, essential workers should receive the vaccine first since they are frequently exposed and can treat other members of the population. The elderly should be amongst the first given the high fatality rates amongst those that contract COVID-19. The vaccine should go to impoverished individuals living in overcrowded and unhealthy conditions, where the disease is mostly likely to spread quickly. Current evidence suggests that members of certain races and ethnicities have been particularly hard hit by COVID-19 [34], suggesting that they, too, should be prioritized. Given the crowded conditions, prisons represent an at-risk population for the spread of disease, therefore prisoners should also be among the first to receive the vaccine (which might represent a difficult political prospect). At the international level, the argument could be made that countries that have done a good job stopping the spread, such as New Zealand, should be the last to receive the vaccine. AI provides an opportunity to bypass complex arguments by focusing explicitly on easily measurable outcomes, while thinking through and including other health outcomes.

The discussion within this article focuses on ethics and policy ideas, considerations, and implications related to the new information created by the AI for the purpose of vaccinating a population with a constrained number of vaccines. We intentionally centre the discussion on new problems and ramifications our technology may create and avoid the discussion around the existing guidelines and general concerns. Our goal is to augment what is already done; we do not wish to discount existing ethics and policy issues discussed in the literature by relevant experts [11], [22].

Despite this focus, we hope that the discussion demonstrates the scale and scope of considerations when applying AI to real-world problems in general. We discuss popular normative ethical lenses through which one may look, and hope they provide a framework for others. Additionally, our goal is not to explicitly apply the ethical theories, but present them as well defined paradigms. In addressing these considerations, we expect the effectiveness of results for real-world problems to be improved, and we hope to make a small contribution towards improving the trust of the general population in AI.

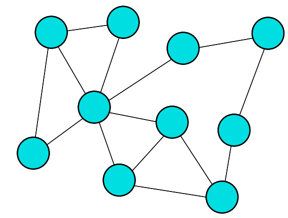

Fig. 1. Small example of a social contact network containing 10 individuals. Nodes represent individuals and edges (connections between nodes) represent connections through which a disease may be transmitted.

II. Simulation & Methods

How best do we apply a limited number of vaccines to a population such that the impact of the disease is minimized? This question is not novel. Many organizations, including the WHO and CDC, have a number of guidelines for addressing this very question; however, they tend to focus on implementation issues, demographic information for specific individuals, and small sub-populations (e.g. front-line workers such as healthcare providers). These guidelines and considerations are critically important and should not be discounted. Our contribution is to include additional information to augment existing practices based on a more *holistic* understanding of the pattern of social contacts within the population.

A social contact network may be represented as a graph $G = (V, E)$ where the vertices (nodes) V represent individuals and the edges E between the vertices represent interactions between individuals that may facilitate the transmission of a given disease. Figure 1 provides an example of an small community of 10 individuals. The topologies of the graphs differ depending on the community and can be impacted by factors such as geography (e.g. rural or urban) and social practices (e.g. implementing social distancing) of a population. The exact social network for a community is rarely known. It may be approximated by public health officials using

[1]We emphasize the importance of understanding *systematic* discrimination and biases. The *systematic* designation describes policies and practices that are part of a larger system that perpetuates disadvantages or biases. An example is the collection and calculation of disease infection rates while assuming a population is homogeneous. Although the average infectious rate may be appropriate for the average individual in the sample population, it may not be representative of minority groups. Although no malice is involved in this example, the problem may be perpetuated throughout the system if the data is used by researchers to build models and make conclusions.

contact tracing, cell phone data, or networks induced from epidemiological data.

In the *SIR* model of infection [31] the population is separated into three mutually-exclusive groups: *Susceptible* (individuals who have yet to become infected), *Infectious* (those currently infected and who can infect others), and *Removed* (those who have recovered or who have succumbed to the disease). This can be expanded to the *SEIR* model [7], with the addition of an *Exposed* group, consisting of individuals who are not yet symptomatic and who can not yet infect others.

At the present time, we use the SEIR model, as it best matches the behaviour of COVID-19 based on current knowledge, with the *Exposed* category being able to represent lengthy incubation periods. It is also been a popular model for studying COVID-19 within the literature.

The transition of an individual from one state to the other is typically governed by probabilistic parameters. Individuals that are susceptible may transition to an exposed state with probability β for each interaction with an infectious individual. Exposed individuals probabilistically transition to the infectious state based on the average incubation/latent period of the disease (α). An infectious individual will transition to removed with some probability γ. These parameters are determined empirically for a given disease within a population.

We are currently developing a more general model called the SEE'IR model. This model includes a new state E' in which individuals are presymptomatic but able to spread the disease. Additionally, individuals may transition from E' to R directly, which represents individuals that are entirely asymptomatic.

Once the parameters are set, a small number of individuals within a given graph are set as infectious and the simulation begins. Each iteration of the simulation represents one time step, which we use to represent a single day.

Without any form of intervention, this simulation simply shows how a given disease will spread throughout a graph loaded into the system. System design allows us to apply mitigation or vaccination rules for the population based on graph properties and other information about the state of the epidemic. A simple strategy could be to vaccinate individuals that have many contacts, or perhaps to vaccinate those that have at least some number of infected neighbours. The discovery of an effective strategy for a given disease and a given graph is non-trivial, and it is here that we use artificial intelligence to automate the process.

The resulting strategies can then be tested within our simulation. In our initial tests, we constrained the system to only apply a small number of vaccines (4% of the population) once every week (7 time steps of the simulation). We also assume that a new *shipment* of vaccines is obtained once per week. The values used are arbitrary and easy to change.

As we operate under the condition that a constrained number of vaccines is available, we wish to identify a strategy that maximizes the number of individuals protected while minimizing the number of vaccines used. Refer to Figure 2 for an idealised scenario in which a single vaccination can protect many individuals within a population. In this example, the infectious individual is in contact with only one person and vaccinating this person will protect the rest of the population regardless of its size.

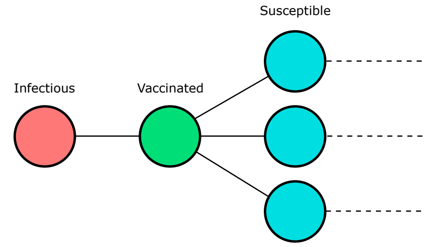

Fig. 2. Ideal scenario in which a single vaccinated individual (green) can stop the spread of a disease from an infected individual (red) to the susceptible individuals (cyan), which form the remainder of the network.

In the SEIR model, we have an additional added complexity — we are unable to detect the difference between susceptible and exposed individuals. Since vaccines are only effective on susceptible individuals, it is important that the strategy does not *waste* vaccines on those that are likely already exposed.

The forms of AI used so far are different forms of evolutionary computation. Although any form of appropriate AI may be used, it is important to focus on *explainable AI*, as discussed in Section IV.

It is critical to understand that the system being developed is not designed to replace any current guidelines, but rather to provide additional information to augment existing practices. Further, although the optimization criteria can easily be changed, the system is still only capable of optimizing what it is explicitly instructed to optimize. The objectives in early experiments [8] included minimizing the maximum number of individuals infected at a given time, maximizing the number of individuals left as susceptible, and minimizing the total number of individuals infected. Note that none of these criteria to date include front-line worker designations, ethical considerations, or demographic information.

III. NORMATIVE ETHICAL FRAMEWORKS

Normative ethical theories make reference to, interact with, and condition moral entities, concepts, and principles. There are several normative approaches one may take when trying to decide what one *ought* to do. Once all (or much of) the information is considered, some of these approaches will converge and recommend an action and other normative approaches might disagree and recommend other actions; however, this tension is useful. We will see that tension often presents itself as a challenge: improve the welfare of as many individuals as possible, while not violating the dignity of any individual in particular. The normative ethical systems are the *heuristics* that aid us in negotiating that tension. Beauchamp and Childress expand considerations of welfare and dignity to include autonomy (consent), beneficence and non-maleficence (two parts of the welfare whole), and justice (fairness, equity, access, information) [9]. Applying the normative theories means using rules that constrain and promote these principles in some programmatic way. Here we describe four broad and influential — and arguably the most dominant and approachable — ethical systems, namely, Consequentialism, Deontological Ethics, Virtue Ethics, and Pragmatism.

A. Consequentialism

Consequentialism places a focus on the outcome (consequence) of a given action/inaction and avoids precepts, edicts, and laws (if they are made free of consideration of consequences — rules and laws that lead to good outcomes may be encouraged). There are many subcategories of consequentialism, but perhaps the most well known is *Utilitarianism*, which focuses on the total *happiness* a given single act (act utilitarianism) or system of rules (rule utilitarianism) creates [40]. With utilitarianism, acts that maximize happiness for the greatest number are considered *good*.

What is meant by "happiness" is problem dependent and, within the context of a single problem, may differ significantly. Happiness is also a rather difficult metric to measure. Measurable proxies may be the economic well-being of a group, increasing total knowledge, the dollar cost/value, or, in our example of vaccination strategies, this may be minimizing the total number of individuals that contract the disease, or perhaps maximizing total health equity.

In the context of welfare, dignity, autonomy, beneficence/ non-maleficence, and justice, a utilitarianism decision rule would take these as inputs and output the action/decision that maximizes the beneficence or minimizes maleficence, possibly at the expense of some autonomy or justice.

B. Deontological Ethics

Deontological, or *duty* ethics has a focus on one's rights and duties. Right action for this normative theory involves acting according to rules generated from a basic belief in the inviolable value of beings (humans). This category includes Kant's *categorical imperatives* (universal moral laws, individuals as ends, and kingdom of ends) [30], *natural rights* [25], [46], and Rawls' *veil of ignorance* — decision makers have enough information about the consequences of their decisions, but do not know specific information about individuals (age, sex, gender, race, etc.) [29].

C. Virtue Ethics

Virtue ethics focuses on developing intrinsic virtues and moral character as opposed to promoting rules regarding the consequences of actions or an individual's rights and duties. There is a focus on defining virtues, how individuals obtain these virtues (inherit vs. cultural) and how the virtues are applied to real world scenarios [5], [21].

D. Pragmatism

Pragmatism argues that ethical and moral *goodness* can change and evolve over time as new information is made available and discovered through experiment. It acknowledges that the number of variables is very large and that moral criteria will be improved by the pursuit of knowledge [17], [18], [27], [28], [47], [48]. This approach emphasizes the inclusion of the broadest set of stakeholders, the best empirical evidence (including deliberation on past choices), and the most inclusive set of normative decision rule structures, into an arena of discourse. Practically speaking, this means inquiry and searching for solutions to problematic situations. *Good* action emerges out of the process of intelligent inquiry [17].

IV. Discussion

We focus our discussion on new issues related to or resulting from our new technology. We avoid the more general discussion to be had about the current guidelines, health equity, vulnerable groups, and the social value of front-line workers [11], [22], [35].

We deliberately chose to use *explainable AI*, since it generates strategies that are easy to understand and communicate. Without intuition as to why a given strategy is *good* (other than empirical evidence), key decision makers are unlikely to use these strategies. Furthermore, explainable and intuitive AI helps improve public trust and buy-in.

One benefit of an AI system is that it is not bound by human biases and preconceived notions; however, this idea is more nuanced. The tool is intentionally designed to consider only factors directly relevant to vaccination impact on transmission. This is done to provide new information to augment existing guidelines that enable a discourse explicitly designed to contend with ethical decision-making. The system itself is acutely *aware* of the biases it uses and those it leaves out. For example, the system is biased towards minimizing vaccine use and maximizing the total number of individuals left uninfected, while explicitly ignoring who is and is not a front line worker. This is then followed by a step in which humans will apply human biases, ethical, and political knowledge to decide which vaccination strategies developed by the AI to use. Note that the AI system does not develop a single strategy, rather it provides an assortment for consideration.

In this example, the explicit moral parameter is to maximize welfare/minimize suffering, which may be in direct tension with competing desirable outcomes, such as public health equity. The actual moral deliberation can only follow once the strategies are tested and understood. Ultimately, this is a living process that can only be evaluated fully *in vivo* through study, discourse, and deliberation.

A. Early Representative Example

Although the system being developed is in its early stages and all results to date are considered by the authors as preliminary, an interesting precursory mitigation strategy generated is presented as it provides context and enables a discussion on how strategies can be interpreted, and the ethical and policy consequences that arise. The strategy presented here proved to be effective in early tests; however, it is emphasized that it is included here for demonstration purposes only and is not necessarily an effective strategy for reducing the spread of SARS-CoV-2 in general.

Algorithm 1 was generated via genetic programming, and it produced high-quality results on initial testing. It is written

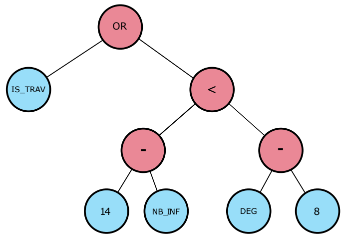

Fig. 3. S-Expression (tree) representation of the preliminary vaccination strategy described in Algorithm 1.

Algorithm 1: Preliminary vaccination strategy. Verbose for the sake of demonstration.
IS_TRAV Boolean indicating if node travels between (connects) communities.
NB_INF number of infected neighbours.
DEG degree of the node.

```
Function Strategy(IS_TRAV, NB_INF, DEG):
    to_vaccinate ← False ;
    if IS_TRAV then
        to_vaccinate ← True ;
    else
        dif_infected ← 14 - NB_INF ;
        dif_degree ← DEG - 8 ;
        if dif_infected < dif_degree then
            to_vaccinate ← True ;
    return to_vaccinate ;
```

verbosely for interpretability. Figure 3 is a simplified and reduced representation of this strategy.

It is critical to remember that one of the purposes and benefits of using AI for generating mitigation strategies is that the AI can *think outside the box*; the AI is not bound by preconceived notions or human biases that may inhibit the algorithm's search[2]. It only optimizes what it is asked to, which, in the case of Algorithm 1, was the number of individuals left uninfected, while also minimizing the number of vaccinations used.

Although we cannot truly know why specific features within the mitigation strategy were important, we can look for intuition within the context of what is being optimized. *Travelers*, for example, are defined as individuals that connect clusters/communities of nodes. Since, in practice, it is difficult to contain community spread, it makes sense to vaccinate these individuals as they facilitate inter-cluster/community spread.

The right sub-tree, as presented in Figure 3, can be further simplified to $NB\_INF > 22 - DEG$, and when considering that the number of neighbours infected (NB_INF) is bound by the degree (DEG) of the node, we find that, with simple algebra, past the value 12, the number of infected neighbours required before an individual is vaccinated decreases as the number of neighbours increases. This also means that any node with fewer than 12 neighbours, regardless of how many infected neighbours they have, is never vaccinated.

[2]Human bias is present when designing the system and interpreting the algorithm's search results — not while the search is being executed.

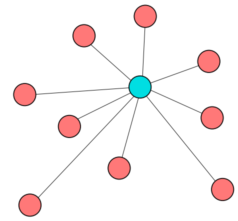

Fig. 4. Susceptible individual (cyan) surrounded by nine infectious (red).

Consider the situation depicted in Figure 4, in which a single, uninfected individual has nine infected neighbours. The generated mitigation strategy would dictate that this individual should *not* be vaccinated. This initially seems rather odd given the circumstances, yet, given more thought, does make sense within the context for which this strategy was being optimized: (a) the SEIR model, (b) minimizing the spread of a disease by maximizing the number of individuals left susceptible, and (c) minimizing the number of vaccinations used.

Given the fact that the *uninfected* individual depicted in Figure 4 has nine infectious neighbours, it is likely that they are not susceptible, but already exposed, and we have yet to identify them as such. If this is the case, then applying a vaccine to this individual would be ineffective, wasting a dose, as within our model vaccines only work on those who are susceptible.

The alternative case is that this individual is still susceptible. If a vaccine is applied in this situation, the vaccine will, at most, protect a single individual. This is not necessarily bad, but remember, the goal is to minimize the spread of the disease throughout the population as a whole; given the constrained number of vaccines, it is ideal for a single vaccine to protect as many individuals as possible (refer to Figure 2).

It is in this latter case where ethical and policy questions, considerations, and implications arise most obviously. The decision of whether the strategy is moral emerges from a reflection and deliberation on the experienced outcomes of the trial. This particular optimization (maximize the number of people not infected while minimizing the total number of vaccines used) strategy leans consequentialist.

Through deontological ethics, it seems harder to defend these strategies, especially when considering natural rights [37] and Kant's second categorical imperative [30]. Kant's imperative demands moral action not be logically inconsistent if made a universal law (does the willing of an act universally result in the destruction of the possibility of acting?), that an action respect our *humanity* as an end not only a means, and that an action leaves the individual as their own law-giver, or generally capable of controlling their own life.

A decision not to vaccinate a particular person, made to ensure the continued operation of the economy, health care institutions, and other systems, might be said to be preserving the necessary conditions for the continuance of one's humanity. Humanity in this sense is a person's set of capacities to pursue their own ends and make their own decisions. This is especially true if that person is given other means to protect themselves under the conditions of full information.

Virtue ethics may also seem to be in tension with such a vaccination strategy, but *may* be consistent. One brand of virtue ethics demands that human flourishing be the guide to right action, and this includes social action. The decision virtue ethics demands must act in accordance with *phronesis* or practical wisdom/excellent judgment. A wise decision is undertaken by an excellent state, and that decision is ethical if it can be shown to promote the virtues (e.g. courage, integrity, prudence, fairness). In certain situations, this would mean *what is best for all*, and in others, it would mean respecting individual autonomy.

Society's responsibility to the individual is to justify actions and help the individual flourish in their particular situation. Has the individual been given the opportunity to flourish, i.e. take courageous, wise, fair, and prudent action (e.g. been informed and resourced to deal with temporary decisions not to vaccinate)?

The pragmatist approach would say that the more information, the better. The AI system was created to demonstrate possible solutions to an emergent problem and was only ever intended to augment the total information available. It is difficult to determine if a given method is moral without the data and discourse. It would also contend that any criticism the AI system receives was only provoked by the action of creating the system and new knowledge; the act of modelling itself is offering data in the pursuit of ethical inquiry. Resulting criticism would be taken and used to modify the system where relevant. It is worth noting that vaccines are not the only tool available for intervention. For example, when considering the case outlined in Figure 4, the individual could be helped by either removing them to a quarantine hostel or mobilizing existing healthcare resources to intervene.

B. Classes of Considerations

The AI does not directly incorporate ethical or political questions, considerations, and implications; the AI is optimizing the criteria it is given, which, *for better or worse*, contains no bias or preconceived notions that a human may have. In other words, the exact same reason that we choose to use the AI may in fact be its largest shortcoming when it comes to the problem of actually implementing the developed strategies within a societal context.

As AI practitioners, we must acknowledge these things and avoid developing information, strategies, and technologies that are doomed to fail as a consequence of *siloing* ourselves.

We identify three classes of ethical and policy questions, considerations, and implications that should be considered: (1) ideas that are identified and can be incorporated into the AI; (2) ideas that are identified and acknowledged, but cannot be incorporated into the AI; (3) ideas that are neither identified nor incorporated into the AI, but arise when specialists are working with results.

Although the identification of these classes is not profound, nor are they unique to the application of AI to various fields — we identify them here as they help facilitate the discussion.

For the first two classes, it is critical that discussions with relevant experts and stakeholders take place to be as exhaustive as possible and reduce the number of items within class three.

We must also acknowledge the existence of the third class, since considerations will be missed, individuals applying the information, strategies, and technologies will notice different considerations, and novel problems may arise as a consequence of the application of the new information, strategies, and technologies.

The inclusion of the third class acknowledges that our specializations are limited and we lack critical relevant information and experience. This pragmatist approach [28] must be used to seek relevant information rather than lazily ignore it. The third class is not unique to the application of AI; it exists in every other field of study. However, it is particularly easy to hide behind within the field of applied AI. When we do this, we perpetuate the idea that AI contributes to the systematic biases within society and fuel mistrust of AI.

C. ***Class 1:*** *Ideas that are Identified and can be Incorporated into the System*

It is critical to include ideas in the system that enable new solutions to be created that also reduce systematic biases, prejudice, and discrimination, and incorporate other ethical and political concerns, not only for the effectiveness of the results, but for public buy-in.

Remember, although it is a *double-edged sword*, one of the purposes of using AI is that the system may discover novel ideas not constrained by human preconceptions. Perhaps by including additional information related to ethical or policy concerns that the AI can optimize, we may obtain interesting and novel ethically-laden ideas on how best to address them.

Currently, the largest issue with our system, as a consequence of the data we currently have available, is that the population is treated as a homogeneous group, all having the same disease parameters (infectious rates, average incubation periods and recovery times). This is in spite of the fact that there are different parameter values for different demographics, such as age, race, sex, socioeconomic conditions, and comorbidity. These groups may also form dense clusters within a social network which will impact how a disease spreads through certain cohorts. An added level of complexity related to this problem is that there may be a poor record of information related to parameter values for certain groups. For example, undocumented individuals in the USA typically have no health coverage [50], [52]. A preliminary technique for multiple classes of individuals has begun testing.

This example is not unique to our situation, as many AI practitioners are at the mercy of those who gathered the data. We rely on those gathering the data to account for as many important features as possible and enforce high data quality standards, but despite best efforts, things will be missed, or, for better or worse, the data will go on to be used for purposes outside the scope of what was originally intended.

This is arguably the most quintessential example of — and excuse for — *passing the buck* in our field, thereby perpetuating systematic discrimination and biases [3]. How is this best to be addressed?

We are fortunate that a better *resolution* of demographic information and relevant disease parameters for SARS-CoV-2 are being gathered and reported. As this data is made available, it will be incorporated into the system to produce a better understanding of the spread of the disease, which will, in turn, produce more effective vaccination strategies.

In many other cases, the data is not available and cannot be incorporated into the system. Despite best efforts, the AIs may be perpetuating the systematic biases. In this scenario, should the work not be done out of fear of contributing to systematic problems? If the resulting discoveries contribute to increasing total happiness, then there is a strong consequentialism argument for conducting the research. Or perhaps we suspect that the actions taken would result in much higher suffering in underrepresented groups, and a society willing to cause such suffering lowers the overall welfare of the society, thereby lowering the total happiness. This may also be problematic from a deontological view, as we are explicitly acknowledging the potential for discrimination in the motivation for doing the research. The pragmatist view would require including as many stakeholders (individuals) as possible at each step, and should any be left out, they ought to be represented in other stages to make up for their unfair exclusion (mobilizing other healthcare resources). The best we can do in this situation is to acknowledge the existence of these problems and be careful to communicate them effectively (class 2, as discussed in the next section).

There are certain groups of individuals who are unable to receive a vaccine (some vaccines are not effective for elderly people; some people are allergic; some people have religious objections, etc.). It is possible to have the AI incorporate this information into the vaccination strategies. This will be done in a way to identify randomly distributed individuals and also clusters/communities of individuals who will be unable to obtain a vaccine. The discussion of the ethics and policy issues arising from this will be presented in the next section (assuming the system effectively incorporates the information). A comparison of results using and not using this information could be useful in determining the impact of anti-vaxxers on public health.

Lastly, and perhaps most obviously, eliminating all errors within the software is critical. During development of our software, two significant errors were identified within the system as a result of misusing or misunderstanding design decisions with third party libraries. Making decisions based on incorrect software can have significant ethical, political, or even legal consequences.

D. ***Class 2:*** *Ideas that are Identified and Acknowledged, but Cannot be Incorporated into the System*

Not all related ethical and political considerations can be incorporated into the system. This may be a result of high complexity, or that encoding is effectively impossible. Regardless, we must acknowledge as many as possible for those that will use the systems or knowledge we create. By keeping the stakeholders as informed as possible we hope to address as many potential issues as possible, such as systematic discrimination. Not only will this improve long term results, but it will also improve trust in the systems and aid public buy-in.

1) Pre Hoc: Some of the considerations we must acknowledge will come before the system is used and are out of our control, beyond making informed decisions about them.

How the social contact network (Figure 1) data is gathered/generated may raise a number of issues. If, for example, the data used to develop the vaccination strategies are entirely synthetic approximations of real communities, then it raises questions about the effectiveness of the models. However, in many jurisdictions, this may be the only viable option.

Cellphone companies or organizations like *Google* through *Google Maps* may infer social contact networks based on tracking information and proximity of individuals. Privacy agreements keep those providers from releasing the information for particular individuals, but they do release aggregated, anonymized data sets with artificial noise added so that no individual person can be identified. These data sets are already being used to help track the COVID-19 pandemic. For example, *Google* has published the *COVID-19 Community Mobility Reports*[4] that show changes in how people move around because of the pandemic. The July 27, 2020 report, for example, shows that, in Canada, people have traveled to parks 145% more often during the pandemic than before the pandemic. To generate the most accurate network for specific communities, data from as many individuals as possible should be included; however, without a form of informed consent, this will violate an individual's autonomy, consent, and welfare. Although there are arguments to be made for the sake of the *greater good*, this would be a very difficult decision politically.

Another possibility is to use opt-in contact tracing apps or data from human contact tracing. This option is infeasible in many places because of a large and growing concern over individual data privacy issues [49]. A number of apps have been made available by governments for the purpose of tracing disease vectors. Much discussion has already taken place about the effectiveness, privacy issues, and ethical concerns of such apps [1], [41]. A similar opt-in style app for the purpose of tracking individuals to generate effective social contact networks may be reasonable, but without significant uptake by the population, the networks may not encompass enough information to be a true representation.

Regardless of how the data is gathered/generated, it must be stored somewhere. Proper data security is required as the data may be sensitive and because it may be possible for bad actors to manipulate the data (for example, change information so that certain groups are negatively affected).

[3] Although we have no control over the resulting discourse, we are capable of acknowledging implications of our decisions and incorporating feedback.

[4] www.google.com/covid19/mobility/

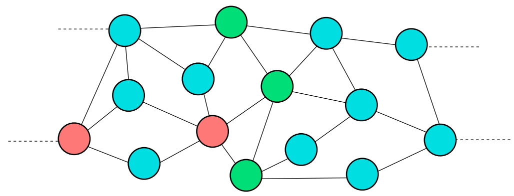

Fig. 5. Walling off a community with vaccines. The right hand community will be protected while the left side may succumb to the disease.

2) Post Hoc: The example provided above in Section IV-A shows that after accounting for many variables, making the "best" decision will depend on the ethical perspective one takes. Despite being difficult to justify from a deontological point of view (counter to natural rights and Kant's second categorical imperative), the decision to not vaccinate the individual in the given example may be a less controversial sell politically with proper public education and discourse.

Public perception and trust are important for the implementation of any strategy, especially those that may appear difficult and unpopular. *Public Health Equity* is declining throughout the world — for example, in the early days of H1N1, the wealthiest nations obtained the vast majority of the vaccine supplies [55]. It is also an increasing public concern [15], [39]. As such, proper and transparent education will be imperative.

It is possible that the system develops an effective strategy that focuses on vaccinating densely populated urban areas. Despite being intuitive, this may be difficult politically as this decision may appear unfair to people living in a rural setting, who already typically have less access to healthcare resources [14]. It may also be the case that particularly remote areas are avoided by the system, which may disproportionately affect Indigenous peoples, a diverse group that has historically been discriminated against and that is continually discriminated against in a healthcare context both directly and systematically [56]. The Canadian territory of Nunavut provides a representative case. It has a small population ($\sim 38,000$) over roughly 2 million km^2, is accessible only by air and sea, is isolated with relative ease, and, as of August 11, 2020, has had zero cases of COVID-19. As a result, AI might completely ignore this population. At the same time, however, Nunavummiut face disproportionately high rates of overcrowding, food insecurity, and respiratory illnesses, and have very limited access to health care and health care facilities, all of which makes this population incredibly susceptible to COVID-19 if it enters the territory.

Since the goal is to incorporate demographic information into the system (as discussed in Section IV-C) we expect to see the AI incorporate this information for the purpose of improving strategy outcomes. However, this will necessarily cause inequalities between groups. There are many examples of AI treating groups differently [57], and even in cases as seemingly benign as social media advertisements, this can have negative consequences for certain groups [4].

Consider the hypothetical scenario in Figure 5 that employs a strategy incorporating network topology and demographic information. Imagine that the left portion is a cluster with higher mortality rates. In this example, the strategy *walled off* this group from the rest of the population, the intuition being that it is too difficult to stop the spread in this group and it is therefore better to simply protect everyone else.

Due to comorbidity complications [19], [54], we may expect strategies prioritizing smokers and obese individuals, groups that already use a large amount of healthcare resources [6]. Also, any solution involving children will create controversy, creating a substantial need for education and discourse. For example, there is evidence that children are less susceptible to the disease [36], but may still spread it. This means that two opposing strategies (do not vaccinate children vs. prioritize only children) may emerge from the AI. Both would be unpopular with the general public.

The above examples are hypothetical, but realistic, and like the example in Section IV-A, put the *greater good* and individual rights and duties in tension. The correct decision for a specific population is impossible to determine at this stage without a broader and inclusive discourse.

Any technology, despite the original intention, may be used for malicious purposes, but the technology itself does not determine human action, nor is it completely free of influence [20]. It is not possible to enumerate all uses of our technology, but we highlight a few interesting cases.

It may be possible for political leaders to specifically target individuals who represent a set of essential voters, or their political base. For example, if a political leader or party typically depends on older rural populations, and remaining in power is their ultimate goal, they would consider it rational to prioritize (reward) this group [16]. Given the network shape and demographic information the technology uses, vaccination strategies may also be weaponized to justify the targeting of clusters of specific at risk minority groups.

A number of decisions are being made that are based on incomplete, current information. As time goes on, our understanding of the disease improves and we obtain more precise information. For example, disease parameters, such as infectious rates and average recovery times are constantly being updated and appear to have different values for different populations. How prevalent are asymptomatic spreaders? Early in the pandemic, it seemed that younger children were low risk; however, there is evidence that acute respiratory distress syndrome (ARDS) related to COVID-19 affects children [45]. For our system, we assume that vaccines are 100% effective, but this may prove to not be the case. We feel that these temporal issues, although unknowns, do not fall into class 3 because we are capable of starting the discourse on the subject.

The system may discover that clusters of individuals not receiving the vaccine may be too debilitating to the greater population. If this is the case, it would be ideal to have mandatory vaccines where possible (an exception would be individuals not able to receive one for health-related reasons). This could come with a number of concerns from the public, as it infringes on individual freedoms, and in a number of countries, it would also be entirely impractical. *Carrot and*

Stick policy has been used in some legal jurisdictions within Canada which require children to be vaccinated in order to attend public school [43], [42]. This provides us with precedents and case studies on the relevant ethical and policy considerations. This is a contentious and long-standing problem placing individual rights and public safety into conflict.

Remember that the current discussion and the new information the technology creates does not happen within a vacuum. With existing information and guidelines and the potential ethical and policy considerations presented here in relation to the new system, it is critical that government decision makers, leaders, and other stakeholders prepare the public effectively. There will be difficult decisions to be made that will be unpopular, but by getting ahead of the problem and educating the public, we may improve outcomes, trust, and uptake.

*E. **Class 3:** Ideas that are Neither Identified Nor Incorporated into the System, but Arise*

The third class exists as an acknowledgement that, despite best efforts, it is inevitable that not all considerations will be in classes one and two. This may arise not from malice, but as a consequence of incomplete information. New ideas will be developed by people using the system as a consequence of having to address novel situations and, perhaps, novel problems becoming apparent as a consequence of the system. The goal is to have as few items in this class as possible when this work is observed by an omniscient individual.

By definition, this list must be empty for the specific problem discussed here since, if anything was known that could be listed, it must be in either class one or two.

V. Conclusions

New technology is currently being developed to discover vaccination strategies based on the structure and shape of social contact networks. These strategies are intended to provide additional information to public health officials and decision makers. The information is not intended to replace any guidelines, but rather to augment them by providing a new dimension of information that takes a more holistic approach.

A number of ethics and policy ideas, considerations, and implications related to the new technology being developed are discussed. The identified issues are broken into three classes: issues that can be incorporated into the system; issues that cannot be incorporated, but are acknowledged; and currently unknown issues that may be discovered later. Normative ethical approaches are also introduced. Despite sometimes being in tension, the ethical approaches are used as a framework to inform ethical decision-making.

By acknowledging as many potential issues as possible, we aim to start the relevant discussions and deliberations through the generation of additional information based on action (the development of new technology); we have no intention of passing the buck.

Although our list cannot be exhaustive, and focuses on issues related to vaccination strategies, we hope that it aids insight and contributes to a broader, general discussion around the ethical and political implications of new information generated by AI and other technologies.

As of July 31, 2020, 26 vaccines are in clinical evaluation and 139 more are in preclinical evaluation [44]. Public trust in the vaccines, vaccination strategies, and the technology used to create them is critical to adoption. With social media misinformation [10], [32], [33], governments and trusted official sources must begin the conversation as soon as possible to prepare the public for potentially difficult decisions.

In policy studies, *framing* refers to how policy issues and information are communicated and categorized to organize meaning and create shared understandings [53]. The effective delivery of the COVID-19 vaccine at the domestic level will be contingent upon open and transparent discourse, deliberation, and communication that creates shared understanding amongst the public about which groups should receive the vaccine first. This communication process should be based on a clear and coherent explanation of the health outcomes that dictated this decision-making — outcomes that AI can help to uncover. “As politicians know only too well but ... scientists often forget, public policy is made of language” [38].

VI. Acknowledgments

This research was supported in part by ResearchNS’ Nova Scotia COVID-19 Health Coalition, consisting of Dalhousie University, Dalhousie University Medical Research Foundation, QEII Health Science Centre Foundation, IWK Foundation, Dartmouth General Hospital Foundation, Nova Scotia Health Authority, and Research Nova Scotia.

This research was also supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

This research was also supported in part by the Heaps Chair Endowment Fund at St. Francis Xavier University through the Dr. H. Stanley & Doreen Alley Heaps Chairship, which promotes the constructive interaction of computer science with liberal arts, such as language and communication, ethics, philosophy, political science and social history.

References

[1] Johannes Abeler, Matthias Bäcker, Ulf Buermeyer, and Hannah Zillessen. Covid-19 contact tracing and data protection can go together. *JMIR mHealth and uHealth*, 8(4):e19359, 2020.

[2] ACIP. Advisory Committee on Immunization Practices Policies and Procedures. Technical report, Centre for Disease Control and Prevention, 2018. https://www.cdc.gov/vaccines/acip/committee/downloads/Policies-Procedures-508.pdf.

[3] Faruque Ahmed. U.S. Advisory Committee on Immunization Practices Handbook for Developing Evidence-based Recommendations. Technical report, Centre for Disease Control and Prevention, 2013. https://www.cdc.gov/vaccines/acip/recs/grade/downloads/handbook.pdf.

[4] M. Ali, R. Sapiezynski, M. Bogen, A. Korolova, A. Mislove, and A. Rieke. Discrimination through optimization: How facebook’s ad delivery can lead to biased outcomes. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–30, 2019.

[5] Karl Ameriks and Desmond M Clarke. *Aristotle: Nicomachean Ethics*. Cambridge University Press, 2000.

[6] Ruopeng An. Health care expenses in relation to obesity and smoking among us adults by gender, race/ethnicity, and age group: 1998–2011. *Public Health*, 129(1):29–36, 2015.

[7] Joan L Aron and Ira B Schwartz. Seasonality and period-doubling bifurcations in an epidemic model. *Journal of theoretical biology*, 110(4):665–679, 1984.

[8] D. Ashlock, E. Shiller, and C. Lee. Comparison of evolved epidemic networks with diffusion characters. In *Proceedings of 2011 IEEE Congress on Evolutionary Computation*, pages 781–788, 2011.

[9] Tom L Beauchamp, James F Childress, et al. *Principles of biomedical ethics*. Oxford University Press, USA, 2001.

[10] Thor Benson. Twitter bots are spreading massive amounts of covid-19 misinformation. *IEEE Spectrum*, July 2020. https://spectrum.ieee.org/tech-talk/telecom/internet/twitter-bots-are-spreading-massive-amounts-of-covid-19-misinformation.

[11] T J Bollyky, L O Gostin, and M A Hamburg. The equitable distribution of covid-19 therapeutics and vaccines. *JAMA*, 2020.

[12] Alex Boyd. When we get a COVID-19 vaccine the real question will be: Who's first? *The Toronto Star*, July 2020. https://www.thestar.com/news/canada/2020/07/28/when-we-get-a-covid-19-vaccine-the-real-question-will-be-who-first.html.

[13] David Burke. How AI and smartwatches can help us cope with COVID-19. *CBC News*, June 2020. https://www.cbc.ca/news/canada/nova-scotia/ai-covid-19-technology-smartwatches-research-1.5588421.

[14] Leighton Chan, L Gary Hart, and David C Goodman. Geographic access to health care for rural medicare beneficiaries. *The Journal of Rural Health*, 22(2):140–146, 2006.

[15] Jason Corburn. Urban place and health equity: critical issues and practices. *International journal of environmental research and public health*, 14(2):117, 2017.

[16] Bruce Bueno De Mesquita and Alastair Smith. *The dictator's handbook: why bad behavior is almost always good politics*. PublicAffairs, 2011.

[17] John Dewey. How we think, a restatement of the relation of reflective thinking to the educative process. *Ankara Üniversitesi İlahiyat Fakültesi Dergisi*, 8, 1933.

[18] John Dewey. *Logic, the theory of inquiry, 1938*. Southern University Press, 1991.

[19] William Dietz and Carlos Santos-Burgoa. Obesity and its implications for covid-19 mortality. *Obesity*, 28(6):1005–1005, 2020.

[20] Andrew Feenberg. The ambivalence of technology. *Sociological Perspectives*, 33(1):35–50, 1990.

[21] Giovanni R Ferrari and Tom Griffith. *Plato: The Republic*. Cambridge University Press Cambridge, 2000.

[22] David P Fidler. Negotiating equitable access to influenza vaccines: global health diplomacy and the controversies surrounding avian influenza h5n1 and pandemic influenza h1n1. In *Negotiating and Navigating Global Health: Case Studies in Global Health Diplomacy*, pages 161–172. World Scientific, 2012.

[23] Ryan Flanagan. Ethical dilemma: Who should get a COVID-19 vaccine first? *CTV News*, June 2020. https://beta.ctvnews.ca/national/coronavirus/2020/6/10/1_4977914.html.

[24] Luciano Floridi, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, et al. Ai4people—an ethical framework for a good ai society: opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4):689–707, 2018.

[25] Thomas Hobbes. Leviathan (1651). *Glasgow 1974*, 1980.

[26] Angus Reid Institute. COVID-19: Three-in-five worry about side-effects of a vaccine; many plan to take a 'wait and see' approach. Technical report, Angus Reid Institute, 2020. http://angusreid.org/coronavirus-vaccine/.

[27] W James. What pragmatism means. lecture 2. *Pragmatism: A new name for some old ways of thinking.–New York*, pages 17–32, 1907.

[28] William James. Pragmatism: a new name for some old ways of thinking (1907). *Cambridge: Harvard UP*, 1975.

[29] Rawls John. *A theory of justice*. Cambridge, Harvard University Press, 1971.

[30] Immanuel Kant. Grounding for the metaphysics of morals (jw ellington, trans.). *Indianapolis, IN: Hackett.(Original work published 1785)*, 1993.

[31] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

[32] Allen Kim. Nearly half of the twitter accounts discussing 'reopening america' may be bots, researchers say. *CNN*, May 2020. https://www.cnn.com/2020/05/22/tech/twitter-bots-trnd/index.html. why-people-spreading-coronavirus-anxiety-on-twitter-might-actually-be-bots-134802.

[33] Ryan Ko. Meet 'sara', 'sharon' and 'mel': why people spreading coronavirus anxiety on twitter might actually be bots. *The Conversation*, April 2020. https://theconversation.com/meet-sara-sharon-and-mel-

[34] Jonathan Kopel, Abhilash Perisetti, Ali Roghani, Muhammad Aziz, Mahesh Gajendran, and Hemant Goyal. Racial and gender-based differences in covid-19. *Frontiers in Public Health*, 8:418, 2020.

[35] Connal Lee. How do we choose who gets the flu vaccine in a pandemic – paramedics, prisoners or the public? *The Conversation*, May 2017. https://theconversation.com/how-do-we-choose-who-gets-the-flu-vaccine-in-a-pandemic-paramedics-prisoners-or-the-public-74164.

[36] Ping-Ing Lee, Ya-Li Hu, Po-Yen Chen, Yhu-Chering Huang, and Po-Ren Hsueh. Are children less susceptible to covid-19? *Journal of Microbiology, Immunology, and Infection*, 2020.

[37] John Locke. Two treatises of government, 1689. *The anthropology of citizenship: A reader*, pages 43–46, 2013.

[38] Giandomenico Majone. *Evidence, argument, and persuasion in the policy process*. Yale University Press, 1989.

[39] Michael Marmot, Sharon Friel, Ruth Bell, Tanja AJ Houweling, Sebastian Taylor, Commission on Social Determinants of Health, et al. Closing the gap in a generation: health equity through action on the social determinants of health. *The lancet*, 372(9650):1661–1669, 2008.

[40] John Stuart Mill. *Utilitarianism*, volume 3. Routledge, 1895.

[41] Jessica Morley, Josh Cowls, Mariarosaria Taddeo, and Luciano Floridi. Ethical guidelines for covid-19 tracing apps, 2020.

[42] Ontario Ministry of Health. Immunization requirements for school attendance: information for health care providers. Technical report, Government of Ontario, 2018. http://www.health.gov.on.ca/en/pro/programs/immunization/ispa.aspx.

[43] Office of the Chief Medical Officer of Health (New Brunswick). Required immunization of school children. Technical report, Government of New Brunswick, 2016. https://www2.gnb.ca/content/dam/gnb/Departments/h-s/pdf/en/CDC/HealthProfessionals/NBIPG-policy_2-9-e.pdf.

[44] World Health Organization. Draft landscape of covid-19 candidate vaccines. Technical report, World Health Organization, 2020. https://www.who.int/publications/m/item/draft-landscape-of-covid-19-candidate-vaccines.

[45] Clare E Pain, Susanna Felsenstein, Gavin Cleary, Sarah Mayell, Karsten Conrad, Srikrishna Harave, Phuoc Duong, Ian Sinha, David Porter, and Christian M Hedrich. Novel paediatric presentation of covid-19 with ards and cytokine storm syndrome without respiratory symptoms. *The Lancet Rheumatology*, 2020.

[46] Thomas Paine. The rights of man (1791). *Collins Dictionary of Quotations*, page 502, 1984.

[47] Charles Sanders Peirce. The fixation of belief. *Philosophy after Darwin: Classic and contemporary readings*, pages 39–48, 1877.

[48] Charles Sanders Peirce. How to make our ideas clear. *The nature of truth: Classic and contemporary perspectives*, 2001:193–209, 1878.

[49] SAS. Data Privacy: Are You Concerned? Technical report, SAS Institute, 2018. https://www.sas.com/en/whitepapers/data-privacy-110027.html.

[50] Gadi Schwartz, Guadalupe Venegas, and Daniel Arkin. California's undocumented food workers face grim dilemma: health or wages. *National Broadcasting Company*, August 2020. https://www.nbcnews.com/news/amp/ncna1236038.

[51] @TorontoStar. *When we get a COVID-19 vaccine the real question will be: Who's first?, Twitter*, July 28, 2020, 11:35 PM EDT (accessed July 28, 2020). https://twitter.com/TorontoStar/status/1288135896177553408.

[52] Aaron van Dorn, Rebecca E Cooney, and Miriam L Sabin. Covid-19 exacerbating inequalities in the us. *Lancet (London, England)*, 395(10232):1243, 2020.

[53] Merlijn Van Hulst and Dvora Yanow. From policy "frames" to "framing" theorizing a more dynamic, political approach. *The American Review of Public Administration*, 46(1):92–112, 2016.

[54] C I Vardavas and K Nikitara. Covid-19 and smoking: A systematic review of the evidence. *Tobacco induced diseases*, 18, 2020.

[55] Jeanne Whalen. Rich nations lock in flu vaccine as poor ones fret. *The Wall Street Journal*, May 2009. https://www.wsj.com/articles/SB124243015022925551.

[56] Lloy Wylie and Stephanie McConkey. Insiders' insight: discrimination against indigenous peoples through the eyes of health care professionals. *Journal of racial and ethnic health disparities*, 6(1):37–45, 2019.

[57] James Zou and Londa Schiebinger. Ai can be sexist and racist—it's time to make it fair, 2018.