The Case for Globalizing Fairness: A Mixed Methods Study on Colonialism, AI, and Health in Africa

Mercy Nyamewaa Asiedu*
Google Research
United States of America
masiedu@google.com

Awa Dieng*
Google DeepMind
Canada
awadieng@google.com

Iskandar Haykel†
Americans for responsible innovation
USA
alexander.haykel@gmail.com

Negar Rostamzadeh
Google Research
not given
nrostamzadeh@google.com

Stephen Pfohl
Google Research
United States of America
spfohl@google.com

Chirag Nagpal
Google Research
United States of America
chiragnagpal@google.com

Maria Nagawa
Duke University
United States of America
maria.nagawa@duke.edu

Abigail Oppong
Ghana NLP
Ghana
abigoppong@gmail.com

Sanmi Koyejo
Google DeepMind
not given
sanmik@google.com

Katherine Heller
Google Research
United States of America
kheller@google.com

Abstract

With growing application of machine learning (ML) technologies in healthcare, there have been calls for developing techniques to understand and mitigate biases these systems may exhibit. Fairness considerations in the development of ML-based solutions for health have particular implications for Africa, which already faces inequitable power imbalances between the Global North and South. This paper seeks to explore fairness for global health, with Africa as a case study. We conduct a scoping review to propose axes of disparities for fairness consideration in the African context and delineate where they may come into play in different ML-enabled medical modalities. We then conduct qualitative research studies with 672 general population study participants and 30 experts in ML, health, and policy focused on Africa to obtain corroborative evidence on the proposed axes of disparities. Our analysis focuses on *colonialism* as the attribute of interest and examines the interplay between artificial intelligence (AI), health, and colonialism. Among the pre-identified attributes, we found that colonial history, country of origin, and national income level were specific axes of disparities that participants believed would cause an AI system to be biased. However, there was also divergence of opinion between experts and general population participants. Whereas experts generally expressed a shared view about the relevance of colonial history for the development and implementation of AI technologies in Africa, the majority of the general population participants surveyed did not think there was a direct link between AI and colonialism. Based on these findings, we provide practical recommendations for developing fairness-aware ML solutions for health in Africa.

CCS Concepts

• **Social and professional topics** → **Medical information policy**; **User characteristics**; • **Computing methodologies** → **Artificial intelligence**; **Machine learning**; • **Human-centered computing** → **Collaborative and social computing**.

Keywords

Algorithmic Fairness, Colonialism, Health, Africa

---

*Authors contributed equally to this work.
†Work done while at Google

1 Introduction

Machine learning (ML) models have the potential for far reaching impact in health. However they also have the potential to propagate biases that reflect real world historical and current inequities and could lead to unintended, harmful outcomes, particularly for vulnerable populations [27, 41, 50, 75, 82, 96].

In recent years, algorithmic fairness literature Barocas et al. [15], has proposed different techniques to understand, evaluate, and mitigate algorithmic biases that may cause harm or contribute to unfair or biased decision making throughout the machine learning development and deployment pipeline. However, these have been mostly contextualized within western notions of discrimination.

Fairness in Global health: Fairness is especially important for global health, which has been plagued with inequitable power imbalances between high-income countries (HICs) and low- and middle-income countries (LMICs) [12, 49]. There are multiple definitions of global health but we use the definition proposed by Beaglehole [17]: *"[Global health is defined as] collaborative trans-national research and action for promoting health for all."* The recent decolonizing global health movement sheds light on "how knowledge generated from HICs define practices and informs thinking to the detriment of knowledge systems in LMICs"[36]. Eichbaum and van Schalkwyk. [36] explore intersections between colonialism, medicine and global health, and how colonialism continues to impact global health programs and partnerships. Given that most machine learning models are developed with problem formulation, resources and data from HICs, and may be imported with little regulation to LMICs, there is a risk for algorithmic colonialism and oppression, detailed extensively by Birhane [20], Mohamed and Isaac [69], which would further exacerbate current power imbalances in global health. On the other hand, increasing internet access and the democratization of machine learning knowledge and tools present an opportunity to create sustainable change in global health by empowering research partnerships to develop ML tools for locally relevant applications in health.

Contextualizing fairness: To date, fairness in health research has furthered understanding and mitigation of biases across the machine learning development pipeline [27, 28, 96]. However most are contextualized to HICs. There are a few studies that have explored contextualizing fairness beyond the West.

Sambasivan and Prabhakaran [90] utilizes mixed methods studies to identify subgroups, proxies and harms for fair ML in the context of India. For natural language processing applications, Bhatt and Prabhakaran [19] identify India-specific axes of disparities such as region and caste, and redefine global axes in an Indian context, such as gender, religion, sexual identity and orientation, and ability. Fletcher and Olubeko [37] provide detailed recommendations for fairness, bias and appropriate ML use in global health with examples using pulmonary disease classification in India.

Philosophical foundations: Recent critical approaches to algorithmic fairness [54, 90], have problematized the philosophical scope of algorithmic fairness research, which thus far has largely been focused on distributive justice, i.e., justice which is concerned with the apportionment of benefits and burdens across members of a society. Such focus, it has been argued, is to the detriment of other important and relevant moral frameworks, including theories of social justice and non-western conceptions of normativity. The result is an inadequate approach to conceptualizing and implementing algorithmic fairness practices. We share this concern. Our contribution in this paper to globalizing fairness practices can be understood as part of an attempt to motivate the development of a

more philosophically inclusive and contextually attuned approach to algorithmic fairness. In certain contexts, such as in certain non-western settings, fairness may require that ML-based interventions take into account factors which matter beyond the narrow question of how to fit such interventions into a fair distributive scheme. As the qualitative research results of our paper suggest, one such factor may be the history and legacy of colonialism.

Contributions: This paper builds on previous work [19, 37, 90] by exploring fairness attributes for global health, with Africa as a case study. Throughout the paper, we will use "fairness attribute" to mean historical, demographic, or social properties (e.g., religion, age, socio-economic status, etc.) which may constitute "axes of disparities" in that they may serve as focal points around which disparities in the performance of ML-based interventions between Africa and the West may be measured. Axes of disparities are factors that run across the entire machine learning pipeline including problem selection and deployment considerations, whereas fairness attributes are a subset of the axes of disparities more specific to the data and model development. We make the following contributions:

- We propose (1) axes of disparities between Africa and the West, and (2) global axes of disparities in the context of Africa, identifying fairness attributes that should be considered with respect to Africa, and delineating real world ML for health applications where these should be considered (section 2).
- We conduct a qualitative study on the relevance of these attributes and on the impact of colonialism on the development and deployment of ML systems in health in Africa (section 3)
- Finally, we discuss contextual challenges to the application of ML in health in Africa and provide recommendations for fairness-aware ML solutions for health in Africa (section 4).

While there are common themes that may resonate across the continent, **Africa is a diverse continent** and there are attributes and axes of disparities which may apply to some countries and not to others.

2 A Scoping Review of Fairness Attributes in Consideration for Africa

In this section, through a scoping literature review, we identify axes of disparities on a global scale, a continental scale, a local scale, and an individual scale (Figure 1). Common axes across most countries include colonial history (90% of sub-Saharan Africa [10]), race (80% fall under the definition of racially black [87]), and national-income levels (85% of countries fall under world bank categories of low income, or lower-middle income [45]). Within the continent there are inter-country disparities, for example national income levels (e.g., 6 countries are upper-middle income and only Seychelles is high income [45]), the North African versus sub-Saharan African divide [18]. There are also intra-country disparities across individual income levels, the rural-urban divide, ethnicity, and ethnic language, which we discuss in detail. We intentionally do not look at one specific "homogeneous" region on the continent in order to highlight unfairness that can affect the continent as a whole, as well as different levels of unfairness at the sub-regional and national levels. These axes of disparities are defined with the African
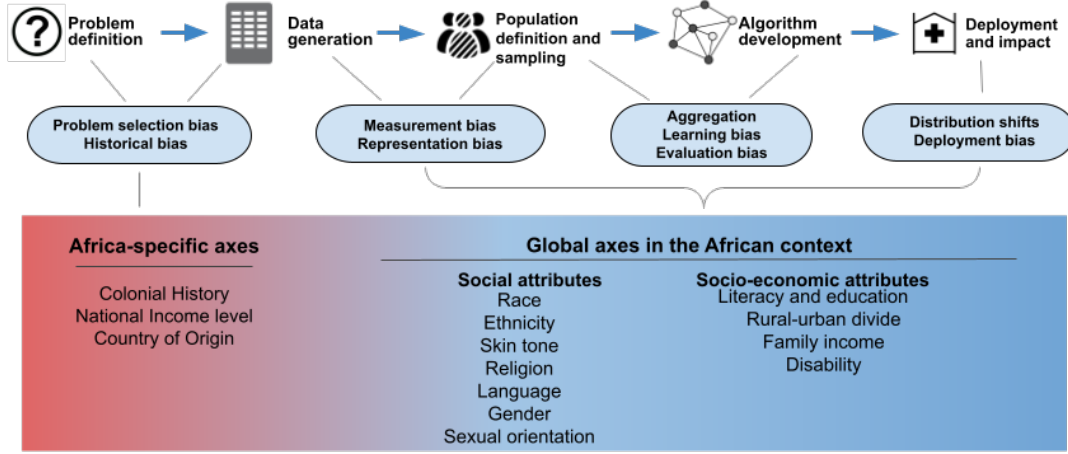
Figure 1: Categorization of African-contextualized axes of disparities with the type of biases they can induce along the machine learning pipeline. ML pipeline and biases figure modified from [28] and [98].

Table 1: Summary of the proposed axes of disparity. They are not mutually exclusive and some are connected

Country-level axes of disparities	Individual-level axes of disparities		
Colonial history[†]	Race[†,§]	Rural-urban location[‡]	Religion[†,§]
National income level (NIL)[†]	Ethnicity[†,§]	Literacy and Education level[‡,§]	Age[‡,§]
Country (origin[‡] and residence[‡])	Language (ethnic)[‡,§]	Skin tone[†,§]	Gender[‡,§]
Language (official)[‡]	Sexual orientation[†,§]	Socio-economic status[‡,§]	Disability status[‡,§]

Legend: [‡]:Historically collected and/or used, [†]: accessible but not historically collected and/or used, [§]: Sensitive.

context in mind, but the underlying logic can be applied to different geographical regions. We define key terms in Appendix A.

2.1 Axes of disparities between Africa and Western Countries

Colonial history: There is strong evidence that Africa's colonial history and resulting structural challenges continue to create power imbalances between the formerly colonized on one hand and former colonial rulers and beneficiaries on the other [36]. Colonial history has been put forward by several scholars as a social determinant of health [30, 85, 102] and this has implications for data collection. We propose colonial history as a fairness attribute that could cause historical, representation, measurement, learning, evaluation, and deployment biases, and one that should be considered throughout the ML pipeline i.e. problem formulation, data collection, model development, and deployment. Learning bias is especially of note as objective measures for model performance should not be limited to accuracy, but should also seek to improve impact metrics.

National income level (NIL): Ranked by NIL, Africa comprises 24 Low-income countries (LIC), 17 lower-middle income countries (LMIC), 6 upper-middle income countries, and 1 high-income country [72]. Socio-economic disparities resulting from the disproportionately high number of LICs and LMICs on the continent imply limited funding from governments for research in ML for health,

and limited availability of clinical data-generating devices and resources for ML development [71]. This results in problem definitions that are more in line with external funding requirements than with local needs[71]. Infrastructure limitations also impact access to data generation medical devices (e.g., mammograms), as well as electronic health storage systems, which can lead to representation bias and limited technical (e.g., compute) and personnel resources for ML development.

Country of origin: As discussed above, different countries have different developmental and national income levels, and varying methods of implementing health strategies [71], which are influenced by each country's leadership and political landscape. Hence machine learning applications that work in one country may not necessarily transfer to the other. Aggregate bias that may result in a one-size fits all model should be avoided by understanding implications for fairness that may be country-specific.

2.2 African-contextualized global axes of disparities

We provide a contextual framing of globally applicable axes of disparities and illustrate why they should be considered both in importing ML models from outside Africa as well as when developing ML models within the continent.

Race: Race in fairness typically refers to structural racism referenced as a social construct due to the history of slavery and

subsequent racial disparities in countries like the United States. Such a construct does not apply directly to Africa. The majority of Africans (80%) are racially black and are subjugated to global anti-black rhetorics [83]. North African Arabs are externally racially white, as are Afrikaans in South Africa. There are also people of Asian origin, predominantly in the east and southern Africa. However, most Africans do not self-identify by race and are more likely to identify by nationality and ethnicity [11, 63]. Demographic health questionnaires do not typically ask for "race," though persons of African descent may be at higher risk for certain diseases.

One may consider racial disparities that impact Africans on a global scale, or within different countries in Africa, between Black and White North Africans [83], or between Black and White South Africans [105]. This has implications for socio-economic disparities in health access, representation bias, and for race-based disparities in ML model performance.

Ethnicity: Africa has over 3000 ethnic groups and is the most genetically diverse continent. Unlike in HICs, most Africans associate with an ethnic group. Ethnicity, in addition to physical traits, language, and culture, defines "historical experience, aspirations, and world view" [32].

Socio-economic disparities in accessing health may run along ethnic lines [38]. This can lead to ML applications skewed towards dominant ethnic groups resulting in ethnic inequities.

Religion: Africans practiced traditional religions pre-colonialism. With the introduction of Western religion, Christianity and Islam have grown to become dominant religions on the continent. Religion can be associated with ethnicity and socio-economic status, leading to "ethno-religious disparities" in health [43, 44]. Religion can also impact access, perceptions, and adoption of health practices [74, 91, 106]. Religion becomes a sensitive attribute when harmful stereotypes and inequities tied with ethnicity are propagated in machine learning models.

Language: There are over 2000 languages in Africa, and most Africans speak more than one [48]. Oral language in Africa typically has the format of a dominant colonial language such as English, French, Arabic, or Portuguese [76].

These dominant languages, particularly colonial languages, facilitate cross-ethnic interaction, are typically both written and oral, are utilized in education and professional life, and have "led to linguistic inequality" [76]. On a per-country level, disparities occur between countries with English as an official language and others due to the "political, economic, and educational...power of the English language" [84].

Skin tone: On a global scale, skin tones of racially black Africans, light or dark, are associated with blackness and instigate anti-blackness discrimination [83].

Among black Africans, lighter skin tone may be perceived as enabling better social status and favors [73]. As a result, 40% of people on the continent practice cosmetic skin bleaching [13]. Bleaching in itself is a public health concern, increasing risk for severe health problems [60]. Within algorithmic development, studies have found algorithmic biases against darker skin tones [22, 33]. In applications to health, this has implications for biases towards lighter skin which may be better represented in training data and cause discrepancies in dermatology applications, such as skin disease detection [31].

Gender: Historically patriarchal African culture, reinforced by colonial legacies, has led to inequitable distribution of wealth between genders [53]. Female household and child caring duties, as well as professional duties more recently, further reduce access to basic health amenities for African women [8]. Transgender and gender non-conforming persons in Africa are stigmatized and discriminated against, with few health programs catering for transgender persons [62]. This also has implications for health access and treatment, especially for sexually transmitted diseases [103].

Sexual orientation: Thirty-two countries criminalize homosexuality and it is punishable by death in 3 countries [59]. Homosexual persons are stigmatized and discriminated against which causes exclusion and marginalization within health systems [88, 94].

Literacy and Education level: Literacy and education levels impact access to care, health seeking behaviors, and understanding of health information, especially when delivered digitally [7]. This has implications not only for one's own health, but also for the health of one's children [23].

Age: Age is a global fairness attribute used both in machine learning model development and evaluations for health, due to age-specific incidence rates and co-morbidity [67].

Rural-urban divide: People living in rural areas may have disproportionate levels of lower socio-economic status, literacy, education, and limited access to health facilities [79, 99]. This makes people in rural areas most vulnerable to unethical machine learning practices.

Socio-economic status: Individual socio-economic disparities underlie most of the above axes of disparities mentioned. Africa has the second highest wealth distribution gap [93]. This runs the risk of machine learning models in health perpetuating unfairness in deployment access or inaccuracies predominantly towards poorer persons.

Disability: 10% to 20% of African populations are affected by disabilities. Disabilities can exacerbate most of the attributes listed due to stigma and inadequate resources and policies at the country level [3, 65].

3 A Qualitative Study on the perceptions of the fairness attributes and the connections between colonialism, health, and AI in Africa

We conduct a mixed methods study with experts and general population participants to study perceptions of the proposed axes of disparities as potential sources of bias in the development and deployment of machine learning models. We focus on colonialism as a central attribute of interest. The study and findings serve to provide evidence for the pre-identified fairness attributes, and how they can be accounted for when building ML-based solutions in health in Africa. We specifically sought to answer the questions 1) *Are disparities along axes of the proposed attributes likely to be sources of bias for ML models?* and 2) *Does colonialism have an impact on the implementation of AI solutions for health in Africa?* In the rest of this section, we describe the study design, participant selection, and findings. To accommodate the diversity of backgrounds in the participants, we used AI and ML interchangeably in the survey and interviews. Refer to Appendix B for questions asked.

Table 2: Demographic distribution of general population and expert survey participants

Demographic	Subgroup	Expert Participants (n=50)	Gen. Pop. Participants (n=672)
Gender	Female	19 (38%)	271 (40.3%)
	Male	30 (60%)	401 (59.7%)
	Non-binary	1 (2%)	0
Age	18-20	0	56 (8.3%)
	21-29	11 (22%)	330 (49.1%)
	30-39	25 (50%)	201 (29.9%)
	40-49	12 (24%)	66 (9.8%)
	50-59	2 (4%)	14 (2.1%)
Country(residence)	South Africa	5 (10%)	128 (19.0%)
	Kenya	8 (16%)	125 (18.6%)
	Nigeria	6 (12%)	169 (25.2%)
	Ghana	3 (6%)	125 (18.6%)
	Rwanda	-	125 (18.6%)
	Other countries in Africa (e.g., Egypt, Tunisia, and others)	17 (34%)	-
	Other countries outside Africa (e.g. USA, Germany)	11 (22%)	-

3.1 Methods

3.1.1 **Expert survey and in-depth interviews (IDIs):** We recruited experts in various related fields using online flier distributions and word-of-mouth. Eligibility criteria were 1) researchers in ML, public health experts, health equity activists, entrepreneurs in digital health in Africa, and policy makers working on or in Africa, 2) 5+ years of experience in the relevant fields of work, 3) ability to read and understand English, 4) no affiliation with the institution conducting the research. Interested participants filled out a screener form, and those who met our eligibility criteria were accepted for the study. Participants digitally signed an informed consent form prior to beginning the study. Fifty-seven experts filled out a survey on Qualtrics in which they answered questions related to their perceptions on the proposed fairness attributes. The survey included definitions of machine learning, AI, fairness and bias.

A subset, twenty-eight of the experts, completed an in-depth interview via online video call in which they were asked about the impact of colonialism on AI in Africa, unique challenges of colonized countries with respect to AI implementation, and social dangers and benefits specific to colonized countries. To analyze the interviews: three researchers independently extracted relevant and interesting phrases related to colonialism and grouped them under codes, then discussed and identified general themes from combining the codes. While we selected colonization as the viewpoint, major themes that emerged under this topic were extracted inductively. The full list of expert survey and interview questions can be found in Appendix B, representative quotes can be found in Appendix C, and demographic distribution of experts can be found in Appendix D.
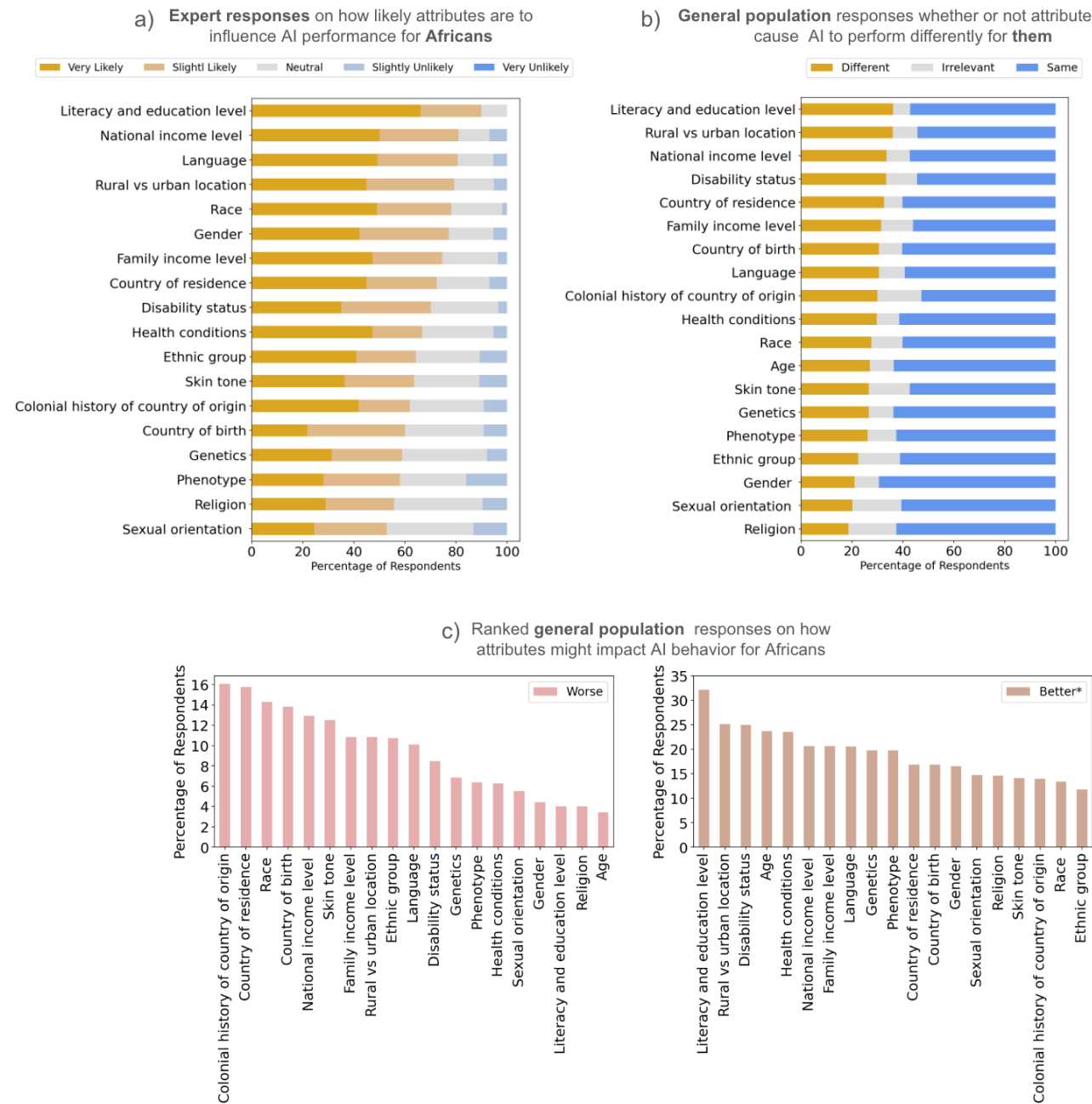
3.1.2 **General population survey.** We surveyed 672 general population participants across 5 countries in Africa (South Africa, Kenya, Nigeria, Ghana, and Rwanda). Surveys were administered by a 3rd party vendor, Gutcheck, and were conducted in English in November 2023. Eligibility criteria were 1) have attained at least a high school diploma or equivalent, 2) ability to read and understand complex and technical text in English "very well," and 3) be at least somewhat familiar with the term "AI" or "ML." Participants were recruited, consented and compensated by the vendor. The survey questions probed participants on the connections between AI and colonialism, and on whether the proposed axes of disparities would lead to disparate performance by AI tools. The survey included definitions of ML, AI, fairness and bias. The full list of questions can be found in Appendix B, representative quotes can be found in Appendix C.

3.2 Results

3.2.1 **Expert and general population opinions on disparities along the axes of the proposed attributes as sources of bias for AI-based tools in health in Africa:**

Most **expert respondents** indicated in a survey that the proposed axes of disparities would impact performance of an AI tool for Africans (Figure 2 a). Each axis mostly (52% - 90%) received responses of being either very likely or extremely likely to cause machine learning models to perform differently for Africans, and none of the proposed axes received an unlikely response from experts. Literacy and education level, national income level, language, rural versus urban location, and race were the top five attributes experts identified as likely to impact AI performance and biases. As one expert put it, *"It [AI] would perform differently for people who do not speak major languages...as their first languages or write it colloquially using their country's version...for men versus women...those who come from ethnic minority tribes particularly marginalised groups...darker skin tones and /or people with different skin from the majority in*

Figure 2: Perceptions of biases for machine learning tools and their associations with African-contextualized by axes of disparities. a) Experts' (n=57) responses on how likely attributes are to influence AI performance for Africans, b) General population responses (n=672) on whether attributes would cause AI to perform differently (better or worse) or the same for them as for others. They could also indicate whether the attribute was irrelevant. c) Breakdown and ranking of general population participant responses on which attributes would perform worse or better for them. *Better is not necessarily good, it reflects that our participant population is better off than other people in the same country. e.g., Most of our survey participants are skewed towards high literacy and education levels compared to the general population, and so AI may perform better for them than for others.

their context... different religious/cultural beliefs from those that align with predominant Western ideas...it would fail to discern between the subtleties across ethnic groups/tribes and countries and the algorithm would prioritise the experiences of countries with large populations like Nigeria and Ethiopia." -P26 [30-39, Nigeria, Public Health]. See appendix C for additional representative quotes.

From the **general population** survey results (Figure 2 b), we found that participants did not generally think AI tools would perform differently for them compared to others, along the proposed axes, which conflicted with expert responses. However, 18% - 36% of respondents indicated that axes of disparities would lead to disparate performance (either better or worse) for them. Only a small fraction indicated that the axes would be irrelevant. Literacy and education level, rural versus urban, national income level, disability status, and country of residence were the five top axes that participants indicated would cause an AI tool to perform differently for them. These were similar to expert results. Colonial history was the most likely to cause an AI tool to perform worse for participants. Literacy and education level were more likely to cause an AI tool to perform better. Given that our general population participants were skewed towards literate and educated backgrounds, this result was not surprising. However, this indicates that people who are not literate or educated could face adverse biases by AI tools. As one respondent indicates, *"AI tools for health in Africa might perform differently across various demographic groups, including urban and rural populations, different age groups, varying socio-economic backgrounds, and distinct cultural communities. Factors like access to technology, healthcare literacy, and local health practices could also influence performance among these diverse groups"* –Gen. Pop participant[21-29, Rwanda]. Sexual orientation and religion were ranked lowest on likelihood for bias for both experts and general population participants. Broken down by country, general population respondents from South Africa were most likely to indicate that an AI tool would perform worse for them, which may be attributed to South Africa being more racially diverse and with black South Africans having most recently been subjugated to apartheid. See appendix C for additional quotes.

3.2.2 The impact of colonialism on the implementation of AI solutions for health in Africa:

Opinions by both experts and general population respondents on the impact of colonialism on today's applications of AI in healthcare are divided. Upon asking experts whether they believed that the history of colonialism impacts the application of AI in former colonies, 16 (57%) said yes, 10 (36%) said no, and 2 (7%) indicated maybe (Figure 3). In the general population survey of 672 people across 5 countries, only 9% thought there was a definite link, 36% thought there might be a link and 55% did not see a link between AI and colonialism (Figure 3). South Africans were most likely to see a definite or slight connection (61%) between AI and colonialism. There was a general consensus between experts on the importance of careful deployment of AI initiatives in order to increase trust and adoption.

Major Theme 1: A prominent theme that emerged was that colonial history has **"kept these countries behind"** and **underdeveloped**, which can potentially lead to the **"lack of ownership"** by the locals, **"economical dependency,"** and **"neocolonialists' exploitation of data and resources,"** which in turn impact **trust**.

One expert participant puts it as *"...[The current] economic situation is a byproduct of colonialism...Ethnic Strife, religious Strife, is also a product of colonialism...that was a tool for keeping people from uniting against colonizers...I'm saying this as somebody whose home country is still colonized, so I have bitter feelings about it"* -P1 [Digital Health, USA, 40-49].

Participants also refer to strong economical dependency to former colonial powers which influences how technologies are adopted in former colonies and how the imbalance of power can lead to AI being developed in only certain sectors that are of interest to former colonialists. Another expert discussing the historical self-interest agendas of the West says, *"historically the British medical system wanted to treat the colonists [and made sure] British military officers were well educated against malaria and tropical diseases and [this education] wasn't meant for the Africans...But when they needed to employ Africans, it was cheaper and Africans were more resistant to African diseases obviously, so that's when it became an issue of interest to the West...so when there are international interventions, they[referring to Nigerians] worry...This was a position [based] solely on an exploitative economic venture by the colonialists''*-P6 [40-49, Nigeria, Medical Doctor]. Thus, some doubt the *"altruism"* that is now demonstrated in global health and AI for social good. P6 continues *"[is it] solidly altruistic, free [data from] blacks, … [or] some economic games, not for Africans?"*. P12 [30-39, Kenya, Machine Learning] adds *"We have the Congo region…there's a lot of political instability…clearly influenced by the colonizers."* Control of African leadership by colonial powers also comes into play, with P28 [30-39, South Africa, Digital Health] saying *"we can't access some resources as fast as other countries, because of the way some of our leaders are still being controlled."* This further influences mistrust *"a few people might think that the West still wants to colonize us using AI. And so this will make an adoption or acceptance of AI products become a challenge"*-P20 [40-49, Kenya, Public Health]. *"...those tools are ways in which our Colonial Masters still have a hold of us"* -P10 [40-49, France, Public health]. There is also apprehension of neocolonialism through leadership - *"There's this AI sector in our country, but it's mostly headed by British and French and it's completely biased towards the way they work and their culture"*-P5 [30-39, USA, Digital Health].

Major Theme 2: Another predominant theme brought up by the participants indicating the impact of colonialism on AI Health application concerns **the existing biases that stem from imbalances in power amongst ethnic, religious and language subgroups due to colonialism and ties to colonial masters.**

For instance disparities in education and health centers by religion is brought up - *"In countries like Malawi and Zambia for example, there is a really strong legacy and heritage of missionaries and their connection to hospitals...so some of the best hospitals were connected to churches"*. Ethnic disparities as a result of colonialist favoritism are also highlighted -*"In Zambia...we have 72 tribes… maybe five or six of them which were sort of favored by the*

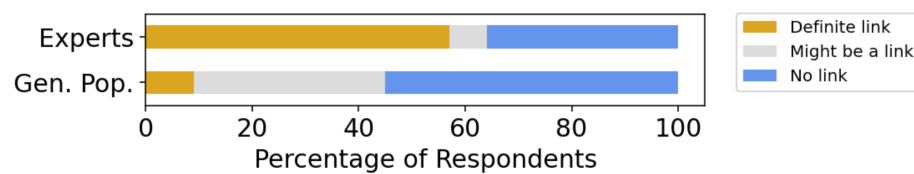**Figure 3: Responses by experts during IDIs (*n* = 30) and general population participants surveys (*n* = 672) on whether they see a connection between AI and colonialism. 57% of experts found a definite link compared to only 9% of general population participants.**

colonial powers….[As such the] western part of those countries are considered generally more educated than the rest of it of the country. So you find that taking technology or taking science of most kinds to the Western North Western southern part of the country is generally easy" -P5 [30-39, USA, Digital Health] and P24 [30-39, USA, Machine Learning] adds that *"a lot of inequalities that exist between different ethnicities in general are a carry-on from colonialism, which is that there are more and more people who… based on their ethnicity actually receive different access to care and that is a consequence of colonialism to a large extent".* The influence of colonial language also can result in disparities in education and technology adoption - *"those who are colonized by the British tend to pick up English…being able to speak English and those colonized by the French tend to have…official language being French, and that influences…research…provision of scholarships [with] preference given to…those colonized by the British"* -P12 [30-39, Kenya, Machine Learning]. P10[40-49, France, Public health] adds to this saying *"technology in general is English-speaking driven…there's a chance that Nigeria, Kenya would have more chances to move forward very fast [than] Senegal and Ivory Coast [which speak French]".*

Other emerging themes on the potentially negative impacts of colonialism on AI:. There were also smaller but relevant themes around Western perceptions of Africa being ingrained into generative models - *"most of the historical data used to build Chat-GPT is the stuff that came from Google Books. It's like books as far back as two thousand years ago. Do I as a woman want to be judged culturally from books from 2000 years ago? No, I do not. "*-P1 [40-49, USA, Digital Health], and around the continued use of frameworks inherited from colonial rule - *"The legal framework the few countries still have some body of laws that are inherited from the colonial past and sometimes these laws have never been touched and are totally outdated."*-P10 [40-49, France, Public health] *"But there are things that have … created… some inequities into the system that we have not been able to fix … after years of colonialism… There are ways in which we practice our medicine… that are based on how the set of doctors who are here, who are trained by whoever the colonial masters were and those behaviors… have seeped into how we still keep practicing medicine now… And so this is the way everybody passes it on to everyone. Now the issue with that is those inadvertently seep into the way patients are treated. They seep into the data."*-P22 [30-39, Nigeria, Medical Doctor]

Major theme 3: There were a few experts who felt there was **no direct impact between AI and colonialism**, or that it was **complicated**.

Some participants talk about forgiveness and the potential of AI to avoid colonial biases if done correctly. *"We do not bring out historical justices…Africans are more forgiving and welcoming…if the launch of AI within Africa is not done correctly, then they will try to associate it with colonial injustice. If done correctly, I don't think Colonial history will be a hindrance."*-P2 [30-39, Kenya, Clinical Officer]. One person indicated that colonialism happened a long time ago and there was a need to move on- *"It's so long ago and we have advanced in our own ways better than before when we were colonized…we are free we can make our own decisions now."* - P20 [40-49, Kenya, Public Health]. Whether or not the respondent's country was colonized impacted responses as well. As P3 [30-39, Ethiopia, Public Health] put it, *"I am living in a country that has never been colonized before, I am from Ethiopia. So I don't think the colonial history of other African countries could have a negative impact on the implementation of AI".* A group of participants indicates the complexity of colonialism and its impact on AI applications, stating that this issue is very context-oriented and can vary across countries. *"[within Nigeria] we don't have a homogeneous perception of how colonialism impacted Nigeria"*-P6 [40-49, Nigeria, Medical Doctor]. *"I would say complex…situations are so different …maybe that the colonial term might be a bit too broad. We just say all African countries were colonized but…it's very different…the application of colonization…Outsiders came in and took resources… then the same outsiders came in and proposed solutions"* -P17 [50-59 Kenya, Machine Learning]. Others attributed limitations in AI not to colonialism but to leadership *"the major problem is maybe leadership…corruption…that is depleting our resources"* -P29 [20-29, Ghana, Machine Learning].

Major Theme 4: Looking forward–towards socially inclusive development of health applications of AI in former colonial spaces.

Despite high associations of AI applications with the history of colonialism and neocolonialism in our expert IDIs, we saw some positive remarks towards making AI socially beneficial for Africa. *"If AI can make Healthcare more accessible, then… it would be a social benefit."* -P2 [30-39, Kenya, Clinical Officer]. Another says *"If for example, all the data in Nigeria is getting into one database…[for] viral infections, you can [attain] reduction in morbidity and mortality by allowing for public health measures"* -P7 [30-39, Nigeria, Machine Learning]. P4 [30-39, Kenya, Policy] mentions colonized countries

may benefit as *"it will be easy to create an awareness on most of the colonial countries in Africa because they are already familiarized with the language in the Western countries such as English and French so that could be a positive impact to implement an AI in colonized African countries."*

There were also forward looking strategies arising in interviews that suggested a careful adoption of AI applications. P24[30-39, USA, Machine Learning] specifically calls for responsible and accessible development of AI- *"if AI is actually built responsibly…[considering] biases and instilled with the community that you're targeting in mind…to actually reach a lot of these communities…[make it] accessible".* And others outline specific methods for preventing colonial tendencies, promoting agency and ownership of AI interventions by historically colonized people, and development of socially beneficial AI for health applications. *"Issues to address include partnerships and networking...working with the local people...the learning institutions to build the trust so that it's not like a tool that you see when you don't work with the key stakeholders right from the start...if it is not so inclusive...[with] the ministries of health, because if they are continuously bypassed...ownership will not be there and then it will now be seen as those are things from the West...So building from scratch together with the communities builds trust, it builds the ownership of the development of that process...that will do away with the decolonizing of healthcare."* -P4 [30-39, Kenya, Policy]. However there were a few participants who saw no path forward towards social benefits from AI for Africa without massive restructuring - *"Nope, not without dismantling and deconstructing the systems and processes by which data is used to feed or build algorithms and predictive models. There are movements of AI for good but when you take a very critical lens...unpack those intentions, the very systems [are built] on structures of oppression and injustice. I think AI for social good would just be a myth"* -P26 [30-39, Nigeria, Public Health].

4 Implications for machine learning development and practical recommendations

In this section, we expand on the themes brought up in the IDI interviews and characterize how they can manifest as unique challenges in the application of ML-based solutions in health in Africa. We end our discussion with practical recommendations for the development and deployment of fairness-aware ML-based solutions in health in Africa.

4.1 Structural challenges to the development of ML-based solutions for health in Africa

In section 3, study participants highlighted the impact of colonialism such that many former colonies are "economically and technologically behind". This has a trickle-down effect on many health applications of ML, which we break down by health modality (Figure 4). Where applicable, we also connect these to the relevant axes of disparities and biases that may result.

Electronic health record (EHR) systems: Currently, machine learning requires digital data for model training or fine tuning. Machine learning for health in HICs have progressed as a result of large amounts of digital health data stored in patient EHRs. In most African countries, the majority of patient health records remain predominantly paper-based due to a lack of infrastructure and computational resources to keep digital copies [77]. This phenomenon is exacerbated in countries with low NIL and in rural areas [4]. Awareness of the crucial role of having EHRs for ML has ushered in new proposals to digitize existing handwritten data and develop culturally appropriate solutions for EHRs.

Medical imaging: Machine learning has shown the most promise for use in medical imaging tasks such as dermatology, pathology, mammography, ultrasound, CT, MRIs and X-rays [26]. However, imaging devices are limited in hospitals in Africa, due to their high cost and maintenance [39, 78]. Even when imaging devices are available, they are of varying quality and interpretation, which leads to poor generalization of ML models from HICs [55, 64, 70]. They are also constrained to urban areas [9, 78], which leads to disparities across the rural-urban divide [55], and are expensive which impacts socio-economic status. Additionally, medical images may not be digitally stored or connected to a patient's other health records due to EHR limitations discussed above. This can lead to limited data for retraining models and can generate representation and evaluation biases especially in instances when the data used to train the model is not representative of African populations, e.g., dermatology classifiers which do not work for darker skin tones.

Lab values: sub-Saharan Africa has the lowest rate of high quality, accredited, clinical laboratories, and while there has been progress through various programs there are still gaps. [5, 40]. Access to laboratory testing services may be limited for people with lower financial means, and those in rural areas without labs who may have to travel to urban-based lab institutions [25, 81]. Given limited EHRs, these results are not always integrated with a patient's comprehensive health profile, leading to segregated health information. In addition to data access limitations, limited infrastructure and personnel (e.g., pathologists) for lab test interpretation impact diagnostic efficiency and accuracy [5, 81]. These present unique challenges for ML purposes which are not observed in HICs.

Survey data: Most application of machine learning to health remain tied to data collected in clinical settings. In the African context, however, there is a breadth of survey data, either self-reported, collected by community health workers, or in clinics as part of initiatives by different non-governmental organizations (NGOs) such as the USAID's demographic health surveys (DHS) [1]. These represent some of the largest, readily available longitudinal sources of health data from the continent and are used to provide data on disease trends over time by different regions and demographics. However surveys especially self-reported ones may be unreliable [6] and may also introduce a number of biases [6, 21], which may impact people in rural areas, those of lower economic status, or those with lower literacy levels [24].

Unstructured written health notes: Machine learning for unstructured health notes utilizes natural language processing models, and more recently large language models to extract information such as symptoms and action items, as well as provide disease classifications [52, 61]. These models are usually pre-trained on large amounts of text data from dominant languages and writing styles, which have been shown to exhibit biases. Using them without proper evaluation and fine-tuning may propagate these identified biases and lead to inaccuracies [80]. There also exist limitations
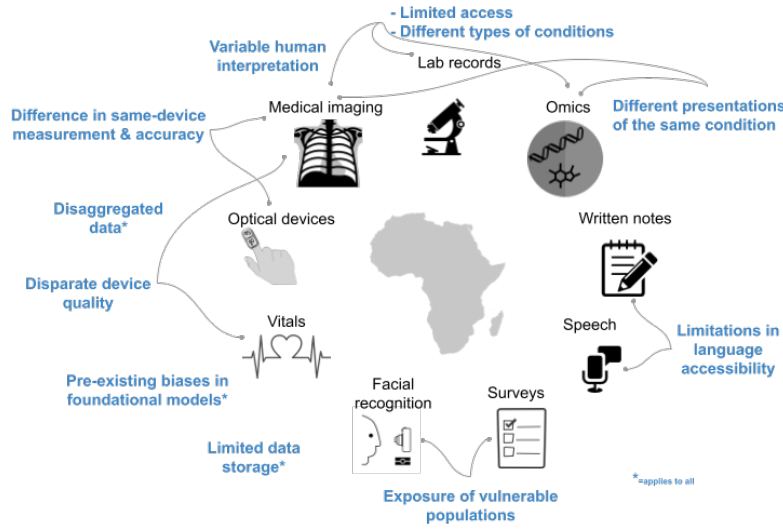
Figure 4: African-contextualized barriers to ML for health by health modality. *=applies to all modalities

to accessing written health notes in health facilities due to limited EHR availability.

Medical speech: Automatic speech recognition systems are used in various heath facilities by healthcare professionals to dictate notes without having to take time away from patient care [101]. Large language models are further driving this application. Accents, style of speaking (e.g., pidgin), and literacy may impact speech recognition algorithms designed to be used in Africa [34, 57, 100]. This can impact language disparities, and may vary by ethnicity, country of origin, literacy, and education.

Optical sensor devices: Optical sensor-based devices such as pulse oximeters and fitness trackers have been shown to have lower performance on darker skinned persons [95]. Machine learning models developed for these devices may perpetuate measurement bias. This has direct implications for skin-tone disparities, which may be indirectly linked to country of origin and ethnicity.

Omics: Using ML to discover biomarkers from multi-omic data is a fast growing field, with potential for precision diagnosis, prognosis and prediction [86]. Africa has large genetic diversity, but available omics data is underrepresented, making up only 1% of omics databases [46]. While the amount of data is increasing, facilitated by consortiums such as Human Heredity and Health in Africa (H3Africa) consortium and the H3Africa Bioinformatics Network (H3ABioNet), there remain limitations in labelled data, ethnic diversity representation, and integration with a patient's health history and profile [46].

4.2 Practical recommendations for fairness-aware ML solutions for health in Africa

Addressing potential machine learning biases starts from structural and systemic changes in healthcare in Africa. (1) Access to basic health care for all persons, (2) increased government and private investment in health access, (3) health-based entrepreneurship and health-related research, (4) sustainable and context-specific

development and deployment of health tools and technologies are essential for reducing disparities in health. Others more specific to machine learning and findings from our qualitative study are:

Contextualization of the fairness criteria: The nuanced differences in axes of disparities between HICs and LMICs lead to a fundamental need to define new or recontextualize existing fairness attributes within the local historical context. There is a unique opportunity to engage different stakeholders (e.g., researchers, policymakers, governance specialists, etc.) to define relevant fairness criteria in accordance with local laws and beliefs. An interdisciplinary approach at this stage is particularly important to avoid potential issues around competing definitions of fairness [29, 56, 104] and purely mathematical formulations that lead to unintended performance degradation for all groups [68].

Collaborative problem selection: As noted in [28] and supported by our qualitative study results, biases stemming from unaligned incentives by external organizations, which lead to power imbalances, can severely affect priority areas, for example which diseases are studied. This has a trickle-down effect in terms of which data is collected regardless of relevance to the local population. It is particularly important at this step to consider the potential biases of stakeholders and ensure that voices of affected communities are included.

Awareness of historical biases towards foreign technology: As our study finds, there is a general mistrust of technologies brought upon by former colonizers (3). Awareness of this existing bias is critical in developing adequate deployment strategies to improve trust and adoption. Without this contextual understanding, deployment of ML-based solutions for health in Africa runs the risk of following historical trends and not taking into account cultural perspectives, under-representation, and other biases.

Caution around using pretrained models: Given the scarcity of readily available large training data sets from Africa, a common ML approach consists of finetuning pretrained models. However, a prevalent concern lies within distribution shifts that may not be

explicit, and the inability to adapt fairness properties to distribution shifts which are likely to occur [92] when deploying models trained on Western-centric data to Africa. In dermatology diagnosis for example, presentations of the same disease are likely to differ between light skin and dark skin tones. Thus models trained on lighter skin may not work when deployed off-the-shelf. Ensuring generalization and transferability of fairness properties under distribution shifts is a new and active area of research [14, 42, 89, 92, 97]. Use-cases specific to the African context can serve as motivating examples and drive impactful advances in this area.

5 Discussion

Initial algorithmic fairness metrics [15] were developed around US legal frameworks of disparate impact and disparate treatment [16]. This centered the work on fairness around contextually relevant protected attributes such as race and gender [51]. We extend these criteria and reframe them in the African context for machine learning applications. Our human-subjects qualitative study amongst general population participants from five African countries, and experts across 17 different countries demonstrates this need and supports the axes of disparities we define. This is in line with previous work on fairness in India that demonstrated the need for recontextualizing fairness criteria [19, 37, 90]. We connect these axes to how they may impact AI for different health modalities and end by providing practical recommendations.

We conduct a deep dive into colonialism as an axe of disparity and while there have been a number of papers on colonialism and AI [2, 20, 69], none of the previous work have conducted qualitative studies to gather expert and African general population opinion. This work demonstrates that experiences may impact Africans' perceptions/opinions about colonialism. People from countries and regions which were not colonized may not associate current challenges faced by health applications of AI with the history of colonialism. Among participants from formerly colonized spaces, there was also a divergence among perspectives with most experts in the IDI indicating a link between AI and colonialism. However in our survey, most of the general population indicated that there was not a definite link. We hypothesize a few reasons for this. Firstly, used in-depth interviews to capture expert responses. This setup enabled probing methods that may have generated a more nuanced, thoughtful responses compared to survey methods used with the general population. Future considerations should consider interviewing a subset of the general population as well. Secondly experts may be have exposure to the topics of decolonizing global health and algorithmic colonization thus care more about AI and colonization compared to the general population.

Recommendations for moving forwards were raised by many of the participants, particularly around accessibility of AI tools, ownership, inclusion, and partnership with local people when it comes to their data. Given the potential connection between AI and colonialism, there is also need for a philosophically broader and contextually attuned approach to algorithmic fairness practices. Such an approach may take into account multiple frameworks for understanding the requirements of fairness–e.g., theories of distributive justice, of social justice, etc. Thus, in the case of Africa, if the history and legacy of colonialism impacts the application of

AI in former colonies, then policy development specialists must work out the implications of this for ensuring the fair development, implementation, and distribution of ML-based interventions.

This work has implications for machine learning models, especially large language and multi-modal models which may perpetuate historical biases and may not have been trained on sufficiently representative data from Africa.

Limitations and future work: There are a number of limitations in this study. Because of the technicality of the survey questions and the need to be able to read and understand English, we exclude a significant amount of the population, especially those who would be most marginalized (low-literacy, rural location). Future work will address this by ensuring better representation, conducting surveys in different languages, and using speech-based methods. Additionally we only conducted IDIs with experts and we will plan to also conduct IDIs with a subset of the general population for more nuanced responses. For future work, we also will explore how to operationalize the proposed attributes in existing bias mitigation techniques. Finally qualitative insights were sought on predetermined axes of disparity, which might limit the discovery of new axes that might emerge from surveys and interviews. However no other axes came out as thematically significant during our analysis.

6 Conclusion

This work lays a foundation for contextualizing fairness beyond western contexts using machine learning applications to health in Africa as a case study. We perform a scoping review to identify axes of disparities in consideration for fairness in Africa and conduct a qualitative study with 672 general population participants and 30 experts participants across several African countries to validate their relevance. Our study also finds that there is a general mistrust of technologies imported from former colonizers among experts, and most associated resource constraints due to pre-existing economic and infrastructure inequities to colonialism. However, the majority of the general population participants surveyed did not think there was a direct link between AI and colonialism. Insights from these findings serve to better understand structural challenges to developing and deploying ML-based solutions for health in Africa as well as provide practical recommendations. We put forth this study as a starting point towards building fair machine learning in global health.

7 Ethical considerations statement

Each participant was provided informed consent and signed a consent form. Participation was completely voluntary and participants could withdraw at anytime. General population participants were recruited and consented through a vendor and were provided an incentive by the vendor, however we did our best in collaboration with the vendor to ensure appropriate incentive range. In-depth interview participants who completed the study were provided an incentive of 100 USD for participating.

8 Researcher positionality statement

All of the study's authors possess qualitative and/or quantitative research backgrounds, and have expertise in data science, artificial

intelligence, machine learning, responsible AI for health and technology policy. The two lead authors are of African origin, as are three of the additional authors. Three of the study's authors are of South Asian and/or Middle Eastern origin. Two of the study's authors are native to the United States. Eight of the authors work on machine learning fairness-related areas in a USA-based corporation. The authors' work is shaped by their lived experiences in Africa and other formerly colonized spaces.

9 Adverse impact statement

We do not anticipate any adverse impacts from this work.

References

[1] 2024. *The Demographic and Health Surveys (DHS) Program.* https://dhsprogram.com/

[2] Rachel Adams. 2021. Can artificial intelligence be decolonized? *Interdisciplinary Science Reviews* 46, 1-2 (2021), 176–197. https://doi.org/10.1080/03080188.2020.1840225 arXiv:https://doi.org/10.1080/03080188.2020.1840225

[3] Fatima Nabbouh-Selvia Shehata Adugna, Molalign B. and Setareh Ghahari. 2020. Barriers and facilitators to healthcare access for children with disabilities in low and middle income sub-Saharan African countries: a scoping review. *BMC health services research* 20 (2020), 1–11.

[4] Maxwell O. Akanbi, Amaka N. Ocheke, Patricia A. Agaba, Comfort A. Daniyam, Emmanuel I. Agaba, Edith N. Okeke, and Christiana O. Ukoli. 2012. Use of Electronic Health Records in sub-Saharan Africa: Progress and challenges. *Journal of Medicine in the Tropics* (2012), 1–6. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4167769/pdf/nihms380682.pdf

[5] Yao K Fonjungo PN. Alemnji GA, Zeh C. 2014. Strengthening national health laboratories in sub-S aharan A frica: a decade of remarkable progress. *Tropical Medicine and International Health* 19.4 (2014), 450–8.

[6] Alaa Althubaiti. 2016. Information bias in health research: definition, pitfalls, and adjustment methods. *Journal of multidisciplinary healthcare* (2016), 211–217.

[7] Padmore Adusei Amoah and David R. Phillips. 2018. Health literacy and health: rethinking the strategies for universal health coverage in Ghana. *Public health* 159 (2018), 40–49.

[8] Seth Kwadjo Angmorterh Dzifa Dordunoo Anaman-Torgbor, Judith and Eric Kwasi Ofori. 2020. Cervical cancer screening behaviours and challenges: a sub-Saharan Africa perspective. *Pan African Medical Journal* 36.1 (2020).

[9] Udunna C. Anazodo, Jinggang J. Ng, Boaz Ehiogu, Johnes Obungoloch, Abiodun Fatade, Henk J. M. M. Mutsaerts, and Mario Forjaz et al. Secca. 2022. A framework for advancing sustainable magnetic resonance imaging access in Africa. *NMR In Biomedicine* (11 2022), 1–19. https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/epdf/10.1002/nbm.4846

[10] Anthony Appiah and Henry Louis Gates. 2010. *C Encyclopedia of Africa.* https://www.oxfordreference.com/display/10.1093/acref/9780195337709.001.0001/acref-9780195337709-e-0467

[11] Godfried Agyeman Asante. 2012. Becoming" Black" in America: Exploring Racial Identity Development of African Immigrants. *Minnesota State University, Mankato* (2012).

[12] Mercy Nyamewaa Asiedu, Awa Dieng, Abigail Oppong, Maria Nagawa, Sanmi Koyejo, and Katherine Heller. 2023. Globalizing fairness attributes in machine learning: A case study on health in africa. *2023 ICLR First Workshop on Machine Learning & Global Health* (2023).

[13] Abubakari Abdulai Edem K. Dzantor Valentine Ayamgba Ayishetu Gariba Grace Buremah-Excellence Asumah, Mubarick N. and Doreen-Remember Donkor. 2022. Prevalence of skin bleaching and its associated factors among young adults in Ghana. *Public Health Toxicology* 2.1 (2022), 1–9.

[14] Ioana Baldini, Dennis Wei, Karthikeyan Natesan Ramamurthy, Moninder Singh, and Mikhail Yurochkin. 2022. Your fairness may vary: Pretrained language model fairness in toxic text classification. In *Findings of the Association for Computational Linguistics: ACL 2022.* Association for Computational Linguistics, Dublin, Ireland, 2245–2262. https://doi.org/10.18653/v1/2022.findings-acl.176

[15] Solon Barocas, Moritz Hardt, and Arvind Narayanan. 2023. *Fairness and machine learning: Limitations and opportunities.* MIT press.

[16] Solon Barocas and Andrew D. Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104, 3 (2016), 671–732. http://www.jstor.org/stable/24758720

[17] Ruth Bonita Beaglehole, Robert. 2010. What is global health? *Global Health Action* 3 (2010).

[18] Ziad Bentahar. 2011. Continental Drift: The disjunction of north and sub-Saharan Africa. *Research in African Literatures* 42.1 (2011), 1–13.

[19] Sunipa Dev Partha Talukdar Shachi Dave Bhatt, Shaily and Vinodkumar Prabhakaran. 2022. Re-contextualizing Fairness in NLP: The Case of India. *arXiv preprint* (2022). https://doi.org/arXiv:2209.12226

[20] Abeba Birhane. 2020. Algorithmic Colonization of Africa. *SCRIPTed* 17.2 (2020), 389. https://doi.org/10.2966/scrip.170220.389

[21] Isakov M Sejdinovic D Meng XL Flaxman S. Bradley VC, Kuriwaki S. 2021. Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature* 600 (2021), 695–700.

[22] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification." In Conference on fairness, accountability and transparency. *PMLR* (2018), 77–91.

[23] Mwoya Byaro and Daniel Mpeta. 2021. Secondary Education and its Effects on Child Health: Empirical Evidence from Sub-Saharan Africa. *African Journal of Economic Review* 9.2 (2021), 116–128.

[24] Tyler Lane Lebo Sello Caroline Kuo Casale, Marisa and Lucie Cluver. 2013. Conducting health survey research in a deep rural South African community: challenges and adaptive strategies. *Health Research Policy and Systems* (2013), 1–11.

[25] Frimpong F.K. Frimpong A.B. Moran Z.R. De Lorenzo M.B. Castañeda-Casado, P. and Y.B. Amor. 2019. Tropical laboratory initiative: an innovative model for laboratory medicine in rural areas. *African Journal of Laboratory Medicine* 8.1 (2019), 1–7.

[26] Leonardo Rundo Marina Codari Giovanni Di Leo Christian Salvatore Matteo Interlenghi Francesca Gallivanone Andrea Cozzi Natascha Claudia D'Amico Castiglioni, Isabella and Francesco Sardanelli. 2021. AI applications to medical images: From machine learning to deep learning. *Physica Medica* 83 (2021), 9–24.

[27] Nigam H. Shah Char, Danton S. and David Magnus. 2018. Implementing machine learning in health care—addressing ethical challenges. *The New England journal of medicine* 378.11 (2018), 981.

[28] Emma Pierson Sherri Rose Shalmali Joshi Kadija Ferryman Chen, Irene Y. and Marzyeh Ghassemi. 2021. Ethical machine learning in healthcare." Annual review of biomedical data science. *The New England journal of medicine* 4 (2021), 123–144.

[29] Alexandra Chouldechova. 2017. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. https://doi.org/10.48550/ARXIV.1703.00056

[30] Karina Czyzewski. 2011. Colonialism as a broader social determinant of health. *International Indigenous Policy Journal* 2.1 (2011).

[31] Kailas Vodrahalli Roberto A. Novoa Melissa Jenkins Weixin Liang Veronica Rotemberg Justin Ko et al. Daneshjou, Roxana. 2022. Disparities in dermatology AI performance on a diverse, curated clinical image set. *Science advances* 8.31 (2022).

[32] Francis M. Deng. 1997. Ethnicity: an African predicament. *The Brookings Review* 15.3 (1997), 28–31.

[33] Wei G Shankar N Thymes B Thorkelsdottir G Kurtz-Miott T Mattson R Obiwumi O Cherepanova V Goldblum M Dooley S, Downing R. 2013. Comparing Human and Machine Bias in Face Recognition. *Preprint: arxiv* (2013).

[34] Atnafu Lambebo Tonja Chris Chinenye Emezue Tobi Olatunji Naome A. Etori Salomey Osei Tosin Adewumi Dossou, Bonaventure FP and Sahib Singh. 2023. Adapting Pretrained ASR Models to Low-resource Clinical Speech using Epistemic Uncertainty-based Data Selection. *arXiv preprint* (2023).

[35] Cynthia Dwork, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. 2012. Fairness through Awareness. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference (ITCS '12).* Association for Computing Machinery, New York, NY, USA, 214–226. https://doi.org/10.1145/2090236.2090255

[36] Lisa V. Adams Jessica Evert Ming-Jung Ho Innocent A. Semali Eichbaum, Quentin G. and Susan C. van Schalkwyk. 2021. Decolonizing global health education: rethinking institutional partnerships and approaches. *Academic Medicine* 96.3 (2021), 329–335.

[37] Audace Nakeshimana Fletcher, Richard Ribón and Olusubomi Olubeko. 2021. Addressing fairness, bias, and appropriate use of artificial intelligence and machine learning in global health. *Frontiers in Artificial Intelligence* 3 (2021).

[38] Raphael Franck and Ilia Rainer. 2012. Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa. *American Political Science Review* 106.2 (2012), 294–325.

[39] Sairam Geethanath and John Thomas Vaughan Jr. 2019. Accessible magnetic resonance imaging: A review. *Journal of Magnetic Resonance Imaging* (06 2019), 65–77. https://onlinelibrary.wiley.com/doi/epdf/10.1002/jmri.26638

[40] Rotz P. Cross D. Belabbes E.H.-Cham F. Ndihokubwayo J.B. Fine G. Zeh C. Njukeng P.A. Mboup S. Gershy-Damet, G.M. and D.E. Sesse. 2010. The World Health Organization African region laboratory accreditation process: improving the quality of laboratory systems in the African region. *American journal of clinical pathology* 134.3 (2010), 393–400.

[41] Suzanne Tamang Jinoos Yazdany Gianfrancesco, Milena A. and Gabriela Schmajuk. 2018. Potential biases in machine learning algorithms using electronic health record data. *JAMA internal medicine* 178.11 (2018), 1544–1547.

[42] Stephen Giguere, Blossom Metevier, Yuriy Brun, Philip S. Thomas, Scott Niekum, and Bruno Castro da Silva. 2022. Fairness Guarantees under Demographic Shift. In *International Conference on Learning Representations.* https://openreview.

net/forum?id=wbPObLm6ueA

[43] Baffour K. Takyi Gyimah, Stephen Obeng and Isaac Addai. 2006. Challenges to the reproductive-health needs of African women: on religion and maternal health utilization in Ghana. *Social science & medicine* 62.12 (2006), 2930–2944.

[44] Peter Salama Stanley Gwavuya Ha, Wei and Chifundo Kanjala. 2014. Is religion the forgotten variable in maternal and child health? Evidence from Zimbabwe. *Social science & medicine* 118 (2014), 80–88.

[45] Catherine van Rompaey Eric Metreau Hamadeh, Nada. 2024. *World Bank Group country classifications by income level for FY24.* https://blogs.worldbank.org/opendata/new-world-bank-group-country-classifications-income-level-fy24

[46] Zass L. Othman H. Radouani F. Allali-I. Hanachi M. Okeke C.J. Chaouch M. Tendwa M.B. Samtal C. Hamdi, Y. and R. Mohamed Sallam. 2021. Human OMICs and computational biology research in Africa: current challenges and prospects. *OMICS: A Journal of Integrative Biology* 25.4 (2021), 213–233.

[47] Moritz Hardt, Eric Price, Eric Price, and Nati Srebro. 2016. Equality of Opportunity in Supervised Learning. In *Advances in Neural Information Processing Systems*, D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett (Eds.), Vol. 29. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2016/file/9d2682367c3935defcb1f9e247a97c0d-Paper.pdf

[48] Bernd Heine. 2000. *African Languages: an Introduction.* Cambridge University Press.

[49] Jens Holst. 2020. Global Health–emergence, hegemonic trends and biomedical reductionism." Globalization and health. *Global Health Action* 16.1 (2020), 1–11.

[50] Galal Galal Mozziyar Etemadi Huang, Jonathan and Mahesh Vaidyanathan. 2022. Evaluation and mitigation of racial bias in clinical machine learning models: Scoping review. *JMIR Medical Informatics* 10.5 (2022), 1544–1547.

[51] Ben Hutchinson and Margaret Mitchell. 2018. 50 Years of Test (Un)fairness: Lessons for Machine Learning. *CoRR* abs/1811.10104 (2018). arXiv:1811.10104 http://arxiv.org/abs/1811.10104

[52] Spasic Irena and Nenadic Goran. 2020. Clinical Text Data in Machine Learning: Systematic Review. *JMIR medical informatics* (2020).

[53] E. O. Jaiyeola. 2020. Patriarchy and colonization: The "brooder house" for gender inequality in Nigeria. *Journal of Research on Women and Gender* 10.1 (2020), 3–22.

[54] Atoosa Kasirzadeh. 2022. Algorithmic Fairness and Structural Injustice: Insights from Feminist Political Philosophy. *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 1–11. https://arxiv.org/pdf/2206.00945.pdf

[55] Kisembo H.N. Remedios D. Malumba R. del Rosario Perez M. Ige T. Hasford F. Brown J.K. Lette M.M. Mansouri B. Kawooya, M.G. and D.H. Salama. 2022. An Africa point of view on quality and safety in imaging. Insights into Imaging. *arXiv preprint* (2022).

[56] Jon M. Kleinberg, Sendhil Mullainathan, and Manish Raghavan. 2016. Inherent Trade-Offs in the Fair Determination of Risk Scores. *CoRR* abs/1609.05807 (2016). arXiv:1609.05807 http://arxiv.org/abs/1609.05807

[57] Allison Koenecke. 2021. Racial Disparities in Automated Speech Recognition. In *Proc. 2021 ISCA Symposium on Security and Privacy in Speech Communication.*

[58] Matt J Kusner, Joshua Loftus, Chris Russell, and Ricardo Silva. 2017. Counterfactual Fairness. In *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett (Eds.), Vol. 30. Curran Associates, Inc. https://proceedings.neurips.cc/paper/2017/file/a486cd07e4ac3d270571622f4f316ec5-Paper.pdf

[59] Kamer Lars. 2022. *Criminalization of homosexual relations in Africa 2020.* https://www.statista.com/statistics/1269999/criminalization-of-same-sex-relations-in-africa/

[60] Karie Gaska Navit Robkin Amber Martin Emily Andrews Lewis, Kelly M. and Justin Williams. 2012. The need for interventions to prevent skin bleaching: A look at Tanzania. *Journal of Black Studies* 43.7 (2012), 787–805.

[61] Irene Li, Jessica Pan, Jeremy Goldwasser, Neha Verma, Wai Pan Wong, Muhammed Yavuz Nuzumlalı, Benjamin Rosand, Yixin Li, Matthew Zhang, David Chang, R. Andrew Taylor, Harlan M. Krumholz, and Dragomir Radev. 2022. Neural Natural Language Processing for unstructured data in electronic health records: A review. *Computer Science Review* 46 (2022), 100511. https://doi.org/10.1016/j.cosrev.2022.100511

[62] Busisiwe Ncama Luvuno, Zamasomi and Gugu Mchunu. 2017. Knowledge, attitudes and practices of health care workers related to treatment and care of transgender patients: A qualitative study in Kwazulu-Natal, South Africa. *Gender and Behaviour* 15.2 (2017), 8694–8706.

[63] Ariela Schachter Maghbouleh, Neda and René D. Flores. 2022. Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. *Proceedings of the National Academy of Sciences* 119.7 (2022), 547–551.

[64] Eric Naab Manson, Francis Hasford, Chris Trauernicht, Taofeeq Abdallah Ige, Stephen Inkoom, Samuel Inyang, Odette Samba, Nadia Khelassi-Toutaoui, Graeme Lazarus, Edem Kwabla Sosu, Mark Pokoo-Aikins, and Magdalena Stoeva. 2023. Africa's readiness for artificial intelligence in clinical radiotherapy delivery: Medical physicists to lead the way. *Physica Medica* 113 (2023), 102653. https://doi.org/10.1016/j.ejmp.2023.102653

[65] Victor McKinney McKinney, Emma L. and Leslie Swartz. 2021. Access to healthcare for people with disabilities in South Africa: Bad at any time, worse during

COVID-19? *South African Family Practice* 63.3 (2021), 116–128.

[66] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. 2021. A Survey on Bias and Fairness in Machine Learning. *ACM Comput. Surv.* 54, 6, Article 115 (2021), 35 pages. https://doi.org/10.1145/3457607

[67] Yuan Zhao Mhasawade, Vishwali and Rumi Chunara. 2021. Machine learning and algorithmic fairness in public and population health. *Nature Machine Intelligence* 3.8 (2021), 116–128.

[68] Brent Mittelstadt, Sandra Wachter, and Chris Russell. 2023. The Unfairness of Fair Machine Learning: Levelling down and strict egalitarianism by default. https://ssrn.com/abstract=4331652

[69] Marie-Therese Png Mohamed, Shakir and William Isaac. 2020. Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence. *Philosophy & Technology* 33 (2020), 659–684.

[70] Daniel J. Mollura, Melissa P. Culp, Erica Pollack, Gillian Battino, John R. Scheel, Victoria L. Mango, Ameena Elahi, Alan Schweitzer, and Farouk Dako. 2020. Artificial Intelligence in Low- and Middle-Income Countries: Innovating Global Health Radiology. *Radiology* (10 2020), 513–520. https://pubs.rsna.org/doi/epdf/10.1148/radiol.2020201434

[71] James Avoka Asamani Nabyonga-Orem, Juliet and Micheal Makanga. 2021. The state of health research governance in Africa: what do we know and how can we improve? *Health Research Policy and Systems* 19 (2021), 1–14.

[72] Catherine van Rompaey Eric Metreau Schweta Grace Eapen Nada, Hamedeh. 2022. *New World Bank country classifications by income level: 2022-2023.* https://blogs.worldbank.org/opendata/new-world-bank-country-classifications-income-level-2022-2023

[73] Wanapa Naravage Nigel F. James Nyoni-Kachambwa, Princess and Marc Van der Putten. 2021. A preliminary study of skin bleaching and factors associated with skin bleaching among women living in Zimbabwe. *African Health Sciences* 11.1 (2021), 132–9.

[74] Phillips E Obasohan. 2014. Religion, ethnicity and contraceptive use among reproductive age women in Nigeria. *International Journal of MCH and AIDS* 3.1 (2014), 63–73.

[75] Brian Powers Christine Vogeli Obermeyer, Ziad and Sendhil Mullainathan. 2019. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 366 (2019), 6464.

[76] M.A Obondo. 2007. Tensions Between English and Mother Tongue Teaching in Post-Colonial Africa. In *International Handbook of English Language Teaching.* Springer International Handbooks of Education.

[77] Florence Femi Odekunle, Raphael Oluseun Odekunle, and Srinivasan Shankar. 2017. Why sub-Saharan Africa lags in electronic health record adoption and possible strategies to increase its adoption in this region. *International Journal of Health Sciences* (10 2017), 59–64. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5654179/pdf/IJHS-11-59.pdf

[78] Godwin Inalegwu Ogbole, Adekunle Olakunle Adeyomoye, Augustina Badu-Peprah, Yaw Mensah, and Donald Amasike Nzeh. 2018. Survey of magnetic resonance imaging availability in West Africa. *Pan African Medical Journal* (07 2018), 1–9. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6295297/pdf/PAMJ-30-240.pdf

[79] Olajide Oloyede. 2017. Rural-urban disparities in health and health care in Africa: cultural competence, lay-beliefs in narratives of diabetes among the rural poor in the eastern cape province of South Africa. *African Sociological Review* 21.2 (2017).

[80] Lester Jenna C. Spichak Simon Rotemberg Veronica Daneshjou Roxana Omiye, Jesutofunmi A. 2023. Large language models propagate race-based medicine. *npj Digital Medicine* 6 (2023), 195.

[81] Christopher R. Polage Thomas C. Quinn Allan R. Ronald Petti, Cathy A. and Merle A. Sande. 2006. Laboratory medicine in Africa: a barrier to effective health care. *Clinical Infectious Diseases* 42.3 (2006), 377–382.

[82] Stephen R Pfohl, Heather Cole-Lewis, Rory Sayres, Darlene Neal, Mercy Asiedu, Awa Dieng, Nenad Tomasev, Qazi Mamunur Rashid, Shekoofeh Azizi, Negar Rostamzadeh, et al. 2024. A toolbox for surfacing health equity harms and biases in large language models. *arXiv preprint arXiv:2403.12025* (2024).

[83] Jemima Pierre. 2013. Race in Africa today: a commentary. *Cultural Anthropol* 28.3 (2013), 547–551.

[84] Asratie Teferra Plonski, Patrick and Rachel Brady. 2013. Why are more African countries adopting English as an official language. *African Studies Association Annual Conference* 23 (2013).

[85] Adriana Garriga-López Ramos, José G. Pérez and Carlos E. Rodríguez-Díaz. 2022. How Is Colonialism a Sociostructural Determinant of Health in Puerto Rico? *AMA journal of ethics* 24.4 (2022), 305–312.

[86] Smarti Reel Ewan Pearson Emanuele Trucco Reel, Parminder S. and Emily Jefferson. 2021. Using machine learning approaches for multi-omics data analysis: A review. *Biotechnology Advances* 49 (2021).

[87] World Population Review. 2024. *How Many Black People are in the World 2024.* https://worldpopulationreview.com/economics/how-many-black-people-are-in-the-world

[88] John Kashiha Lourence Misedah Lucy R. Mgopa Markus Larsson Anette Agardh Ross, Michael W. and Krishna Kumar Venkitachalam. 2021. Addressing the

Healthcare Needs of African Men who have Sex with Men: Barriers to Healthcare and Promoting HIV and STI Treatment in Sub-Saharan Africa. *East African Journal of Health and Science* 3.1 (2021), 59–77.

[89] Bashir Sadeghi and Vishnu Naresh Boddeti. 2020. Imparting Fairness to Pre-Trained Biased Representations. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 75–82. https://doi.org/10.1109/CVPRW50498.2020.00016

[90] Erin Arnesen Ben Hutchinson Tulsee Doshi Sambasivan, Nithya and Vinodkumar Prabhakaran. 2021. Re-imagining algorithmic fairness in india and beyond. *In Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (2021), 315–328.

[91] Thomas E. Olivier J. & Cochrane J. R. Schmid, B. 2008. The contribution of religious entities to health sub-Saharan Africa. *OpenUCT* (2008).

[92] Jessica Schrouff, Natalie Harris, Oluwasanmi O Koyejo, Ibrahim Alabdulmohsin, Eva Schnider, Krista Opsahl-Ong, Alexander Brown, Subhrajit Roy, Diana Mincu, Chrsitina Chen, Awa Dieng, Yuan Liu, Vivek Natarajan, Alan Karthikesalingam, Katherine A Heller, Silvia Chiappa, and Alexander D'Amour. 2022. Diagnosing failures of fairness transfer across distribution shift in real-world medical settings. In *Advances in Neural Information Processing Systems*, Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho (Eds.). https://openreview.net/forum?id=K-A4tDJ6HHf

[93] Okanda J. Seery, E. and M Lawson. 2019. A tale of two continents: fighting inequality in Africa. *Oxfam* (2019). https://doi.org/10.21201/201.4719

[94] Kate Jolly Sekoni, Adekemi Oluwayemisi and Nicola Kay Gale. 2022. Hidden healthcare populations: using intersectionality to theorise the experiences of LGBT+ people in Nigeria, Africa. *Global Public Health* 17.1 (2022), 134–149.

[95] Dumville J Hill J Norman G Hamer O Clegg A Watkins CL Georgiou G Hodkinson A Lightbody CE Shi C, Goodall M. 2022. The accuracy of pulse oximetry in measuring oxygen saturation by levels of skin pigmentation: a systematic review and meta-analysis. *BMC medicine* 20 (2022).

[96] Haytham Siala and Yichuan Wang. 2022. SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. *Social Science & Medicine* (2022), 114782.

[97] Harvineet Singh, Rina Singh, Vishwali Mhasawade, and Rumi Chunara. 2021. Fairness Violations and Mitigation under Covariate Shift. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. Association for Computing Machinery, New York, NY, USA, 3–13. https://doi.org/10.1145/3442188.3445865

[98] Harini Suresh and John Guttag. 2021. Understanding potential sources of harm throughout the machine learning life cycle. *MIT Case Studies in Social and Ethical Responsibilities of Computing* 8 (2021).

[99] T Sven. 2015. Rural health inequities: data and decisions. *Lancet* (2015).

[100] Rachael Tatman and Conner Kasten. 2017. Effects of Talker Dialect, Gender & Race on Accuracy of Bing Speech and YouTube Automatic Captions. In *Interspeech*. https://api.semanticscholar.org/CorpusID:38523141

[101] Mangu Ramya Tai-Seale Ming Lafata Jennifer Elston Tran, Brian D and Kai Zheng. 2023. Automatic speech recognition performance for digital scribes: a performance comparison between general-purpose and specialized models tuned for patient-clinician conversations. *Annual Symposium proceedings. AMIA Symposium* 6 (2023), 195.

[102] Meredeth Turshen. 1977. The impact of colonialism on health and health services in Tanzania. *International Journal of Health Services* 7.1 (1977), 7–35.

[103] Allanise Cloete van der Merwe, L. Leigh-Ann and Donald Skinner. 2022. How HIV advocacy can be used to ensure quality transgender health care: lessons from South Africa. *The Lancet* 399.10330 (2022), 1099–1101.

[104] Sahil Verma and Julia Rubin. 2018. Fairness Definitions Explained. In *Proceedings of the International Workshop on Software Fairness (FairWare '18)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3194770.3194776

[105] C. Verwey and M. Quayle. 2012. Whiteness, racism, and Afrikaner identity in post-apartheid South Africa. *African Affairs* 111 (2012), 551–575.

[106] Peter White. 2015. The concept of diseases and health care in African traditional religion in Ghana. *HTS: Theological Studies* 71.3 (2015), 1–7.

A Definitions

We define terms that are used in this paper. While there is not a standardized definition of fairness, we mainly follow the taxonomy described in Mehrabi et al. [66]. We also provide other terms of general informational value for further understanding of fairness.

Table 3: Axes of Disparity Definitions

Africa-specific axes of Disparities	Disparities that primarily impact Africa.
Global axes of disparities in the context of Africa	Disparities that have global implications but contextualized for Africa.

Table 4: Bias-related Definitions

Discrimination	'Discrimination can be considered as a source for unfairness that is due to human prejudice and stereotyping based on the sensitive attributes, which may happen intentionally or unintentionally'.
Bias	'Bias can be considered as a source for unfairness that is due to the data collection, sampling, and measurement.'
Historical Bias	'Historical bias arises even if data is perfectly measured and sampled, if the world as it is or was leads to a model that produces harmful outcomes. Such a system, even if it reflects the world accurately, can still inflict harm on a population. Considerations of historical bias often involve evaluating the representational harm (such as reinforcing a stereotype) to a particular group'.
Representation Bias	'Representation bias occurs when the development sample underrepresents some part of the population, and subsequently fails to generalize well for a subset of the use population'.
Measurement Bias	'Measurement bias occurs when choosing, collecting, or computing features and labels to use in a prediction problem. Typically, a feature or label is a proxy (a concrete measurement) chosen to approximate some construct (an idea or concept) that is not directly encoded or observable. Proxies become problematic when proxy's are an oversimplification, are measured differently across groups, or accuracy differs across groups'.
Aggregation Bias	'Aggregation bias arises when a one-size-fits-all model is used for data in which there are underlying groups or types of examples that should be considered differently.A particular data set might represent people or groups with different backgrounds, cultures, or norms, and a given variable can mean something quite different across them. Aggregation bias can lead to a model that is not optimal for any group, or a model that is fit to the dominant population'.
Learning Bias	'Learning bias arises when modeling choices amplify performance disparities across different examples in the data. Issues can arise when prioritizing one objective (e.g., overall accuracy) damages another (e.g., disparate impact)'.
Evaluation Bias	'Evaluation bias occurs when the benchmark data used for a particular task does not represent the use '.
Deployment Bias	'Deployment bias arises when there is a mismatch between the problem a model is intended to solve and the way in which it is actually used'.

Table 5: Fairness-related definitions

Fairness	'In the context of decision-making, fairness is the absence of any prejudice or favoritism toward an individual or group based on their inherent or acquired characteristics'.
Individual Fairness	[35]: 'Individual fairness is captured by the principle that any two individuals who are similar with respect to a particular task should be classified similarly'.
Counterfactual Fairness	[58] 'The counterfactual fairness definition is based on the "intuition that a decision is fair towards an individual if it is the same in both the actual world and a counterfactual world where the individual belonged to a different demographic group.'
Group Fairness	Broadly, group fairness notions aim to 'treat different groups equally'. Below are two main statistical group fairness definitions that exist in the literature.
Equal Opportunity	[47]: 'This fairness notion requires that the probability of a person in a positive class being assigned to a positive outcome to be equal for both protected and unprotected (female and male) group members. In other words, the equal opportunity definition states that the protected and unprotected groups should have equal true positive rates'.
Demographic Parity	[35]: ' requires that the overall proportion of individuals in a protected group predicted as positive (or negative) to be the same as that of the overall population'.
Fairness in Relational Domains	'A notion of fairness that is able to capture the relational structure in a domain—not only by taking attributes of individuals into consideration but by taking into account the social, organizational, and other connections between individuals'.
Subgroup Fairness	Subgroup fairness intends to obtain the best properties of the group and individual notions of fairness. It picks a group fairness constraint like equalizing false positive and asks whether this constraint holds over a large collection of subgroups.

B   Survey and interview questions

Table 6: General population survey questions

Q1. Do you see any connection between AI and colonialism? (*Options: Yes/Maybe/No*)

Q2. Do you believe that an automated AI tool would be biased towards or against you? In other words, do you have any reason to believe that an AI-based tool would treat you better or worse than other people? (*Options: It would treat me better than others, It would treat me worse than others, It would treat me the same as others*)

Q3. AI models may perform differently across different groups of people for a variety of reasons. How do you believe an AI tool for health would perform for you and based on the following criteria? (*See axes of disparities criteria and answer options in Table 7*)

Q4. Still thinking about how AI models may perform differently across different groups of people or different locations, what types of groups of people do you believe an AI tool for health in Africa would perform differently for? Please think any group of people, not necessarily a group you may consider yourself a part of. You may refer to the criteria above.(*essay response*)

Table 7: Set up for axes of disparity selections by general population (Reference Table

Rows/ axes of disparities

The country where you were born

The country where you live

Colonial history of your country of origin

National income level (i.e., economic performance of your country) Race (whether you are Black, White, Asian, Mixed race, etc)

Ethnic group you belong to

Your age

Your skin tone

The religion you practice

The language you speak

Your gender (male, female, transgender, non-binary, etc)

Your sexual orientation (attracted to people of opposite gender, attracted to people of the same gender as yourself, attracted to people of any gender, not attracted to anyone, etc)

Your literacy and education level

Where you live - eg. rural or urban location

You or your family income level

A disability - eg. physical disability, blindness, etc.

Any health conditions you currently have (eg. high blood pressure, diabetes, asthma, or others)

Your genetics

Your phenotype (i.e, how your body looks, including shape, weight, height, etc compared to others)

Columns/ Single selection options

AI will perform the same for me as for others

AI will perform better for me than for others

AI will perform worse for me than for others

I do not believe this is relevant

Table 8: Expert survey questions

Experts Survey questions
Q1. Can you please categorize the following criteria by how likely they are to cause an AI tool for health to perform differently (be biased) for groups or sub-groups of people in Africa? (*please refer to Figure5*) Q2. Still thinking about how AI models may perform differently across different groups of people or different locations, what types of groups of people do you believe an AI tool for health in Africa would perform differently for? Please think any group of people, not necessarily a group you may consider yourself a part of. You may refer to the criteria above.

Table 9: Experts In-depth interview questions

Experts In-depth interview questions
Q1. With respect to colonialism, is the application of AI to the health domain in African countries and former colonial territories impacted by the history of colonialism? Q2. Continuing to think about the impact of colonialism, what do you perceive as unique challenges in the application of AI to the health domain faced by African countries and former colonial territories? Probe: Are there specific challenges you can think of related to colonial history? Q3. With respect to colonialism, do you perceive any social benefits involved in the application of AI to African countries and former colonial spaces? If yes, what are some of the social benefits you can think of? Q4. With respect to colonialism, do you perceive any social dangers involved in the application of AI to African countries and former colonial spaces? If yes, what are some of the social dangers you can think of?

Q20

Artificial Intelligence (AI) - based tools may perform differently across different groups of people for a variety of reasons. Can you please categorize the following criteria by how likely they are to cause an AI tool for health to perform differently (be biased) for groups or sub-groups of people in Africa?

	Very Likely	Slightly Likely	Neutral	Slightly Unlikely	Very Unlikely
Country of birth	○	○	○	○	○
Country of residence	○	○	○	○	○
Colonial history of country of origin	○	○	○	○	○
National income level (i.e economic performance of your country)	○	○	○	○	○
Race (whether you are Black, White, Asian, Mixed race, etc)	○	○	○	○	○
Ethnic group	○	○	○	○	○
Skin tone	○	○	○	○	○
Religion	○	○	○	○	○
Language	○	○	○	○	○
Gender (male, female, transgender, non-binary, etc)	○	○	○	○	○
Sexual orientation (attracted to people of opposite gender, attracted to people of the same gender as yourself, attracted to people of any gender, not attracted to anyone, etc	○	○	○	○	○
Literacy and education level	○	○	○	○	○
Rural vs urban location	○	○	○	○	○
Family income level	○	○	○	○	○
Disability	○	○	○	○	○
Health conditions (eg. high blood pressure, diabetes, asthma, or others)	○	○	○	○	○
Genetics	○	○	○	○	○
Phenotype (eg. body look, shape, weight, height, etc compared to others)	○	○	○	○	○

Figure 5: Example of the questionnaire in the survey

C   Representative quotes from surveys on attributes likely to lead to bias

Table 10: Representative quotes from general population

Representative quotes from general population:
"It may perform differently for Africans generally, but even worse for Africans living in the remote areas whose lifestyle is not yet understood lest alone have AI offer correct outcomes or diagnosis." **-Ghana, 30-39, Man**
"AI tools for health in Africa might perform differently across various demographic groups, including urban and rural populations, different age groups, varying socio-economic backgrounds, and distinct cultural communities. Factors like access to technology, healthcare literacy, and local health practices could also influence performance among these diverse groups." **-Rwanda, 21-29, Woman**
In general, there are a number of groups of people who might be disproportionately impacted by an AI tool for health in Africa. For example, people with low literacy levels or those who speak a minority language might have difficulty using the tool, and people living in rural areas might not have access to the necessary infrastructure to use it. People with certain disabilities or medical conditions might also be disproportionately impacted. Additionally, there could be gender or ethnic disparities in access to or use of the tool." **-Nigeria, 30-39, Man**

Table 11: Representative quotes from Experts

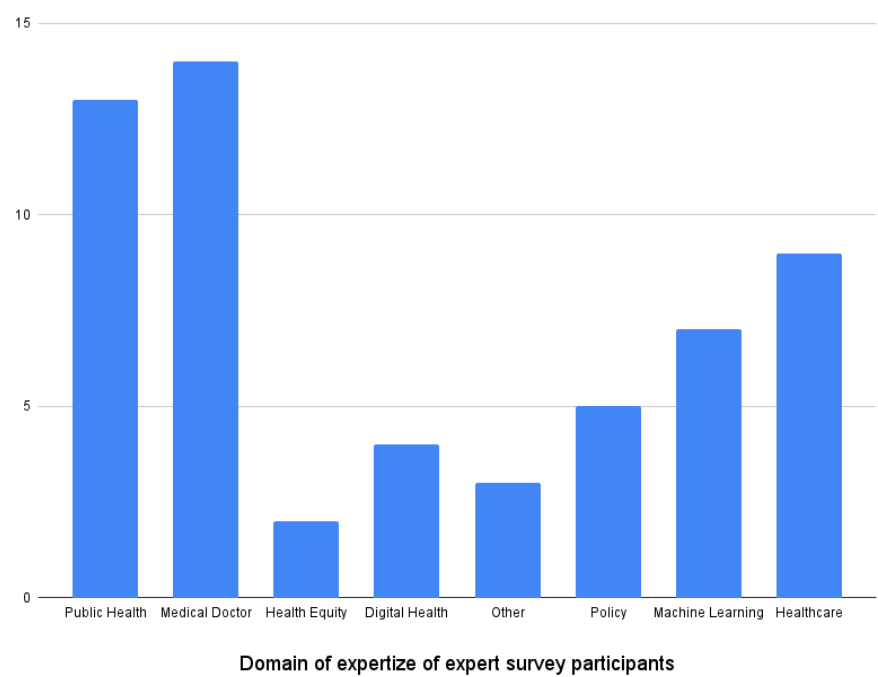| Representative quotes from Experts |
| --- |
| *"I think it would perform differently for people who speak/write do not speak major languages- English, French, Portuguese, or Spanish as their first languages or write it colloquially using their countries version (so a localised version). I think it would perform differently for men versus women, those who come from ethnic minority tribes particularly marginalised groups. I think that it would work very differently for people with darker skin tones and /or people with different skin from the majority in their context...it would perform differently for people who have different religious/cultural beliefs from those that align with predominant western ideas...it would fail to discern between the subtleties across ethnic groups/tribes and countries and the algorithm would would prioritise the experiences of countries with large populations like Nigeria and Ethiopia." -***Public Health researcher, Female, Nigeria, 30-39*** |
| *"The poor and vulnerable would benefit a great deal from AI in healthcare but it's performance might not be great now. African society is religious and the rise of AI is causing agitation and some feel it's an insult to their beliefs. Having AI in health care will meet resistance in religious communities and affect uptake." -***Clinical Officer, Kenya, non-binary, 30-39*** |
| *People with albinism, Pregnant women, People with physical disability, People with mental health disorders, Migrants or displaced people or refugees, People in prisons." -***Policy maker, male,Malawi, 40-49*** |

D   Expert participant distributions

Figure 6: Expert area of expert survey participants

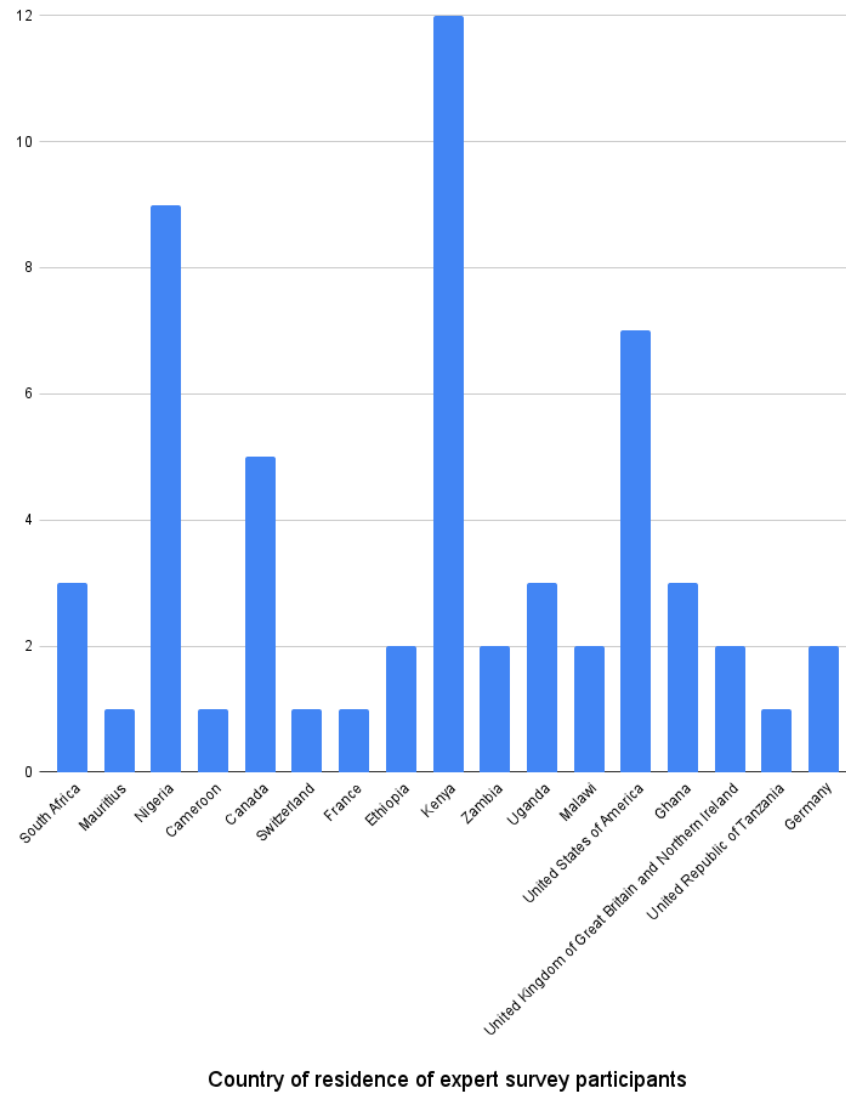Country of residence of expert survey participants

Figure 7: Country of residence of expert survey participants
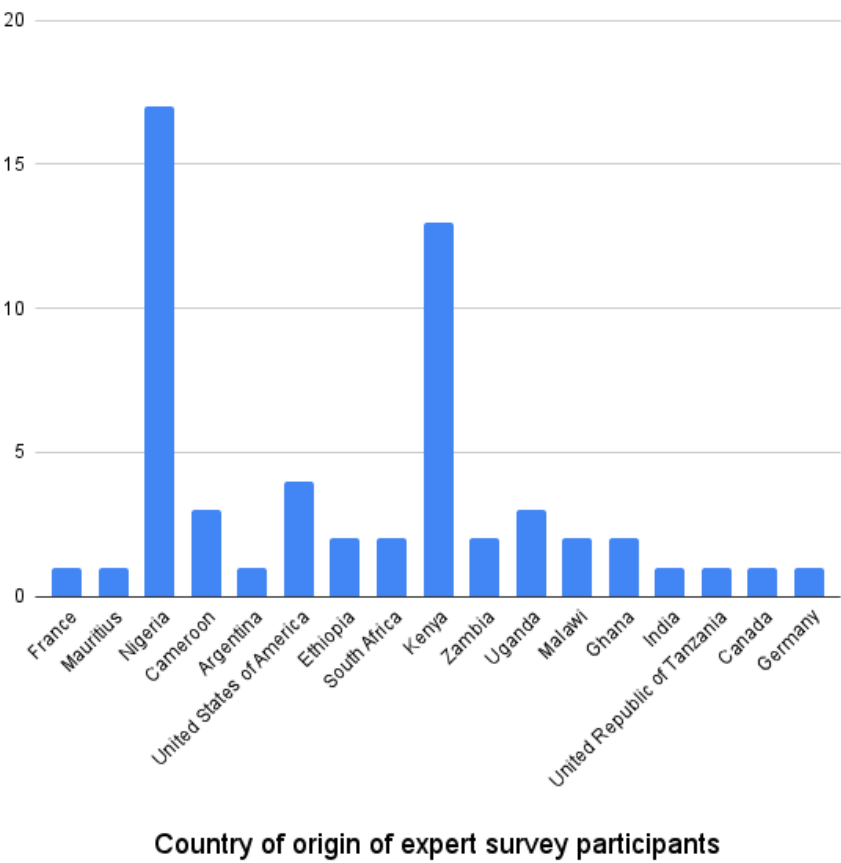
Figure 8: Country of origin of expert survey participants

E Cover Picture showing a word cloud generated from responses from General Population survey when asked about axes of disparities
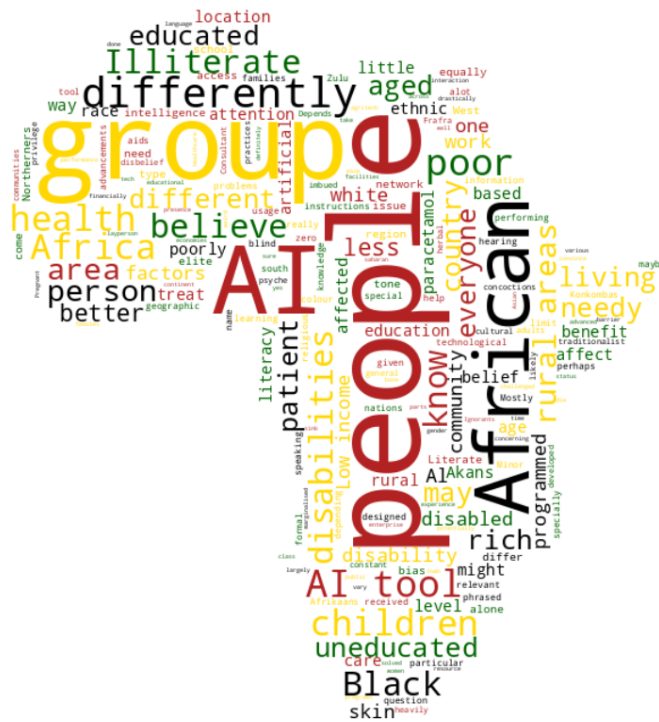

Figure 9: Word cloud of general population verbose responses to the question on axes of disparities