Supplementary information

Supplementary Information A.

NTDs are illnesses prevalent in tropical and sub-tropical areas of the world, where they affect communities living in poor sanitation conditions. Compared to other diseases, investment in scientific research, epidemiological control, and health systems to combat NTDs from 2008 to 2023 remained stagnant or, worse, very low[1,2]. The occurrence and burden of these diseases reflect the vulnerability of the healthcare system caused by the economic crisis in affected countries[3]. NTDs affect more than 1 billion people worldwide and are composed of 21 diseases (Table S1) caused by bacteria, fungi, virus, toxins and parasites[4–6]. These diseases are prevalent in low-and-middle-income countries (LMICs), where the most affected communities live in extreme poverty (survive with less than US$1.25 per day), in war zones, or without any socioeconomic assistance[7,8]. Some of these diseases currently have no treatment, and the cure for others depends on the repurposing of drugs (usually presenting serious side effects and facing multiple drugs resistance). Because of all these factors, NTDs cause in 2019, the loss of 12,048,918 disability-adjusted life-years (DALY) from all ages in East Africa corresponding to US$ 21,824,211,076 and an average of US$1811 per DALY[9].

Table S1. List of NTDs and their prevalent regions

No.	Disease	Etiologic Agent(s)	Class	Geographic Distribution
1	Buruli ulcer	*Mycobacterium ulcerans*	Bacteria	West/Central Africa, Latin America, Asia, Oceania
2	Trachoma	*Chlamydia trachomatis*		Africa, Asia, Latin America, Oceania, Middle East
3	Leprosy	*Mycobacterium leprae*		Southeast Asia, Central Africa, Latin America
4	Yaws	*Treponema pallidum pertenue*		Latin America, Central/West Africa, Southeast Asia, Oceania
5	Chikungunya	Chikungunya virus (CHIKV)	Virus	Latin America
6	Dengue	Dengue virus (DENV)		
7	Zika	Zika virus (ZIKV)		
8	Mycetoma	Actinomycetoma species Eumycetoma species	Bacteria/Fungi	Sub-Saharan Africa
9	Chromoblastomycosis	*Fonsecaea* spp., *Cladophialophora* spp.	Fungi	South/Central America, Africa, Asia, Oceania
10	Chagas disease	*Trypanosoma cruzi*	Protozoa	Latin America
11	African trypanosomiasis (Sleeping sickness)	*T. brucei gambiense*; *T. brucei rhodesiense*		Sub-Saharan Africa
12	Leishmaniasis	Visceral: *Leishmania donovani, L. infantum* Cutaneous: *L. major, L. tropica, L. braziliensis*, others		Sub-Saharan Africa, South Asia, Latin America
13	Schistosomiasis	Intestinal: *S. mansoni, S. japonicum, S. mekongi, S. guineensis*	Parasite	Africa, Middle East, Caribbean, South America, Asia

Continued on next page

Table S1 – continued from previous page

No.	Disease	Etiologic Agent(s)	Class	Geographic Distribution
		Urogenital: *S. haematobium*		Africa, Middle East, Corsica
14	Lymphatic filariasis (Elephantiasis)	*Wuchereria bancrofti, Brugia malayi, B. timori*		Sub-Saharan Africa
15	Onchocerciasis (River blindness)	*Onchocerca volvulus*		Sub-Saharan Africa, Latin America, Middle East
16	Soil-transmitted helminthiases	Various helminths		Sub-Saharan Africa, Southeast Asia
17	Foodborne trematodiases	*Clonorchis sinensis, Fasciola* spp., *Opisthorchis* spp.		Africa, Asia, Latin America
18	Guinea worm disease (Dracunculiasis)	*Dracunculus medinensis*		South Asia, Africa
19	Taeniasis/cysticercosis	*Taenia solium*		Africa, Latin America, Asia
20	Echinococcosis	*Echinococcus granulosus, E. multilocularis, E. vogeli*		East Africa, Latin America, Asia, North America, Europe
21	Scabies	*Sarcoptes scabiei* var. *hominis*	Parasite	Africa, Americas, Asia, Oceania

A.1 Rural vs. Urban Disparities in Healthcare Access

Over 1.7 billion people around the world are afflicted by NTDs, and two-fifths of these people are in Africa[10]. Africa, despite being a mineral-rich continent, has many of its countries that still have large populations living below the poverty line[11]. NTDs primarily affect poor communities in rural areas where access to healthcare, clean water, and sanitation are limited. These communities are usually overcrowded and located in far-flung areas[10,12].

All NTDs are likely treatable and/or curable; they can be contained to prevent disease scourges[12]. However, these diseases continue to be the major causes of morbidity and mortality in affected regions. Accessing healthcare in the far-flung rural communities is met with multiple barriers. These include a shortage of medical facilities, and where such facilities are available, there is often a lack of qualified healthcare professionals, who are usually based in major urban centers[10,13].

Health inequality has been identified as one of the key drivers of the NTD plight. Based on the NTD programme conducted in Cameroon for the Mass Distribution of Drugs (MDA) against Schistosomiasis, a troubling inequality in attaining coverage for the entire population was noted. Immigrant farmers, Women of child-bearing age, and school-aged children were continuously being excluded from the MDA program[13,14].

In the rural areas, the remote locations of some communities are a major barrier to accessing healthcare, as reaching these places is often difficult for medical services, especially during emergencies. Patients also have to travel long journeys to access healthcare. Additionally, some of the medications for treating NTDs are not provided for free; patients have to pay out of pocket[10,13,15].

Between the 19th and 20th centuries, the global urban population grew from 746 million to about 4 billion. Considering the fast growth of urbanization, NTD programs need to be prioritized in these areas[16]. In urban areas, the major barrier to healthcare access is the lack of clear urban governance. The obscurity in the responsibilities between district and municipal authorities poses a challenge in the implementation of MDA programs[17].

A.2 Common Traits of NTDs

Some NTDs are known for their fingerprinted characteristics associated with the clinical manifestation of the disease. Overall, the characteristics are sorted into two groups related to either the physical ability of individuals or socio-economic class[18].

Lymphatic filariasis (Elephantiasis) is a parasitic disease that causes permanent disability, disfiguring swelling (lymphedema) of the limbs, breasts, and genitals (hydrocele)[19]. Leprosy (Hansen's disease) is a disease caused by a bacterium that first affects the skin and peripheral nerves. Without clinical treatment, it can progress and permanently damage the nerves, skin, limbs, and eyes, leading to physical deformities and significant disability[20]. The black-flies–borne disease, onchocerciasis (River Blindness), caused by a parasitic worm, is known for its severe skin problems, intense itching, skin discoloration, impaired vision, and irreversible blindness[21]. The bacterial infection Buruli ulcer results, if not treated promptly, in extensive and destructive skin ulcers by necrosis, responsible for permanent disfigurement and long-term physical disabilities[22].

These diseases, among others not depicted here, often result in intense social stigma and exclusion with devastating psychological impacts (family instability, mental distress, suicidal thoughts, social isolation, and others)[23].

Concerning socio-economic features, NTDs are known to be prevalent in the poorest communities (lack of sanitation and clean water, proximity to infected animals, consumption of contaminated foods) in tropical and subtropical countries[24]. These communities usually do not have political voices and consideration, consequently receiving low resources that could improve daily life[25]. In addition, social and institutional discrimination affects the citizenship rights and participation of people carrying sequelae and scars from NTDs[26]. The negative impact on the economy is devastating, as stigma and other forms of discrimination can lead to school disruption and prevent people in these contexts from working, leading to continuous economic crisis[27].

A.3 Global Prevalence, Incidence, and Disability-Adjusted Life Years (DALYs)

In 2023, an estimated 1.495 billion individuals required interventions for NTDs, representing a reduction of 122 million from 2022 and a 32% decline relative to the 2010 baseline. Between 2015 and 2021, the global burden of NTDs decreased from 17.2 million to 14.1 million DALYs, while annual fatalities linked to NTDs declined from approximately 139,000 to 119,000[28]. The population affected by NTDs fell from 1.9 billion in 1990 to slightly over 1 billion in 2021.

In 2023, 867.1 million individuals received treatment for at least one NTD, with 99% receiving preventive chemotherapy. In 2024, the WHO recognized seven countries for eliminating an NTD: Brazil and Timor-Leste for lymphatic filariasis; Chad for the Gambiense form of human African trypanosomiasis; India, Pakistan, and Vietnam for trachoma; and Jordan for leprosy. These recognitions increased the cumulative total to 54 countries that have eliminated at least one NTD, accounting for 75 successfully completed acknowledgment processes[2].

A.4 Direct Impact

If left untreated, long-term symptoms of the NTDs may lead to chronic diseases, complications or disabilities, from cognitive impairment to infertility and immobility. Lymphatic filariasis causes elephantiasis[29,30], schistosomiasis causes anaemia[31,32] and growth stunting[33,34], while trachoma and onchocerciasis lead to visual impairment and blindness[35–38]. Some NTDs are also associated with cancer and neurological diseases[39], including neurological impairment resulting from rabies, cysticercosis, leprosy and dengue[40]. These complications may lead to chronic disorders and life-long disabilities; for instance, blindness due to trachoma or onchocerciasis limits a person's ability to perform certain tasks independently[41]. Genital schistosomiasis may cause infertility in girls and women[42].

In addition to these long-term consequences, NTDs can have serious implications for maternal health. Maternal parasitic infection increases the risks of anaemia, premature birth and stillbirths[43,44]. Although NTDs are not always associated with death[41], they may be fatal if untreated and contribute to mortality in affected populations, with an estimated 170,000 deaths reported annually[45]. These direct effects highlight the substantial health burden imposed by NTDs on individuals and communities.

A.5 Indirect Impact

Beyond their direct health impacts, NTDs incur significant indirect costs that affect the economy of affected regions. People infected with NTDs are often unable to work, leading to reduced income and household financial burden, with insufficient funds for food, education and healthcare[46]. Consequently, agricultural productivity may decrease as infected individuals are unable to participate fully in farming, contributing to food insecurity[47]. Complications associated with NTDs may require costly interventions, further stressing national healthcare budgets, while malnutrition linked to schistosomiasis may reduce immunity and increase vulnerability to other diseases.

These economic effects extend to the national level. Increases in leprosy, lymphatic filariasis, schistosomiasis and onchocerciasis prevalence reduced the economies of selected African countries by 0.43%, 0.24%, 0.28% and 0.36%, respectively[48]. As a result, NTDs worsen the poverty cycle, as they mostly affect the poor and contribute to deeper poverty[39].

The indirect consequences of NTDs also extend to education. They contribute to school dropouts, leading to high illiteracy rates, low employability and reduced economic productivity[49]. Children may experience absenteeism, cognitive impairment and poor academic performance, while some may become caregivers for affected family members[39], preventing them from attending school.

In addition to economic and educational consequences, NTD-related stigma and discrimination negatively affect family relationships, employment and social engagement, and may delay timely treatment. This stigma is often more pronounced in women[49]. Affected individuals may experience social isolation and mental disorders[50], including depression and anxiety[51]. Fear and denial associated with NTDs may further hinder treatment, and political leaders may deprioritize NTD interventions[49]. Together, these indirect effects underscore the wide-ranging social and economic burden of NTDs.

Supplementary Information B.

Although some success has been achieved to date in the drive to control NTDs following the 2012 London Declaration on NTDs by pharmaceutical companies, donors, endemic countries, and non-governmental organizations[52,53], elimination (interruption of transmission) and eradication targets as envisaged in the United Nations 2030 sustainable development goals (SDGs)[53], particularly SDG 3 (Good Health and Well-Being), currently face some challenges.

To this end, of the 21 NTDs listed in the 2021–2030 WHO roadmap, 16 are presently diagnosed with immunological point-of-care testing tools, 7 with nucleic acid-based assays, and 6, which are skin-related NTDs, are without commercially available tests. This latter group includes leprosy, chromoblastomycosis, mycetoma, taeniasis, dracunculiasis, and rhodesiense human African trypanosomiasis[54,55]. Although other skin-related NTDs—viz. cutaneous leishmaniasis and Buruli ulcer have point-of-care tests that one would ordinarily assume are easily translatable to those without, retrieving diagnostic samples from patients is generally invasive, requiring sterile environments. This highlights the clinical need of non-invasive (e.g., skin tissue surface scrapings) point-of-care tools for improved accessibility and patient comfort[55].

In addition, some current point-of-care diagnostics for NTDs suffer shortcomings; for example, dengue, chikungunya, and lymphatic filariasis have cross-reactivity with other infectious diseases[56], whereas immunological-based point-of-care tools for schistosomiasis, rabies, lymphatic filariasis, Chagas disease, chikungunya, dengue, yaws, and trachoma are inadequate[57]. Nevertheless, despite reports on novel diagnostic techniques for different NTDs based on the detection of parasite secretions in blood, urine, or faeces[58,59], translation of these outputs into clinical application in resource-limited settings at an affordable cost presents a major shortcoming.

Moreover, some NTDs, such as leprosy, Chagas disease, and onchocerciasis, have a long incubation period, thus presenting a challenge in evaluating the relationship between observed cases and underlying transmission, which is useful in disease management and total eradication[53].

Solitary implementation of healthcare interventions within endemic regions with limited and/or poor infrastructure, fragmented data systems, and poor surveillance hinder effective decision-making, often culminating in wasteful and unproductive use of resources and progress monitoring in successfully controlling NTDs[60]. Such has been the case in countries such as Equatorial Guinea, Egypt, Mexico, and India, where lack of funds, capacity, and resources led to resurgence or a delay in disease elimination; Maldives, Ecuador, Mexico, Guatemala, and Bangladesh, where effective interventions were not in place; and Egypt, Maldives, Mexico, Ecuador, Guatemala, Mexico, Bangladesh, and India, which showed insufficient intensity of interventions[61].

Although the WHO has published data on the average annual global number of individuals requiring drug administration for at least one preventative chemotherapy-related NTD, as well as new cases needing treatment for other NTDs[62], these facts remain theoretical statistics that encounter practical limitations. For example, in regions that have very high infection transmission rates or where adult infection is most widespread, the estimated target coverages become too low to prevent transmission[63]. In addition, different age-infection profiles of parasites are differently affected by treatment[53], in turn requiring individualized therapy as opposed to mass drug administration. Furthermore, co-infection is prevalent with most NTDs, requiring unified approaches in terms of diagnosis, treatment, and control[53]. To this end, intensified post-treatment surveillance for residual infection and confirmation of elimination are cardinal in ensuring the attainment of NTD eradication by 2030[64].

B.1 Diagnostic Gaps

Despite progress since the 2012 London Declaration, substantial diagnostic gaps remain that impede timely case detection, patient management, and interruption of transmission for many NTDs. Of the 21 NTDs listed in the WHO 2021–2030 roadmap, a substantial fraction rely on immunological point-of-care (POC) tests (16 diseases) or nucleic-acid based assays (7 diseases), while several skin-related NTDs remain without commercially available diagnostics (e.g., leprosy, chromoblastomycosis,

mycetoma, taeniasis, dracunculiasis, and rhodesiense human African trypanosomiasis)[54,55]. Even where POC tests exist (for example, some tests for cutaneous leishmaniasis and Buruli ulcer), practical deployment is constrained by invasive sample collection procedures that require sterile environments, reducing accessibility and patient acceptability; this underlines the clinical need for truly non-invasive, field-adapted sampling approaches (e.g., surface skin scrapings or urine/blood biomarkers) and their translation into low-cost POC devices[55,58,65].

Current POC technologies also suffer specificity and sensitivity limitations in many settings. Cross-reactivity among arboviral infections (dengue, chikungunya) and between those and other pathogens undermines the diagnostic accuracy of rapid immunoassays, and several high-burden conditions (schistosomiasis, rabies, lymphatic filariasis, Chagas disease, chikungunya, dengue, yaws, and trachoma) continue to be underserved by reliable immunological POC tools[56,57]. Novel biomarker and molecular detections (parasite secretions in blood, urine or faeces) show promise in research settings, but the translation of these methods into affordable, robust tests suitable for resource-limited, high-transmission environments remains limited[58,65]. Finally, long incubation periods for diseases such as leprosy, Chagas disease and onchocerciasis complicate the interpretation of observed cases versus ongoing transmission and reduce the effectiveness of diagnosis-based surveillance unless coupled with sensitive tests and longitudinal follow-up[53].

B.2 Treatment Barriers

The barriers that hindered the control of NTDs can be summarized in different factors such as systematic and operational barriers, lack of awareness campaigns, environmental factors, and internal and external socio-political destabilization. These obstacles include limited funding that subsequently causes a low number of health system infrastructures, lack of structures and facilities, limited logistical access to remote communities, and limited actions that promote academic and technological research. It is noteworthy to mention that poverty, conflicts, and geographical areas are some of the characteristics found in regions devastated by neglected diseases[66].

Low- and middle-income countries geographically located in tropical and subtropical areas are economically disadvantaged in addressing socio-economic problems. Consequently, they usually underfund research on drug development, epidemiological programs, and technology that can improve living conditions. For example, in Brazil from 2004 to 2020, a study revealed that funding for NTD research was stagnated[67]. Apart from that, the weakness of health systems is also due to the occurrence of many endemic areas under the responsibility of a limited number of trained staff, making it difficult to deliver consistent care[68].

The use of communication means (educational campaigns) to alert people is one of the effective methods for preventing NTDs transmitted by mammalian and insect vectors[69,70]. Lack of communication can also cause social problems related to beliefs, misconceptions about diseases, and non-adherence to treatment[26,71]. Environmental factors marked by climate change continue to cause devastating impacts due to flooding, which can spread infectious pathogens quickly and expand endemic areas. Besides, flooding also creates a perfect ecosystem for vectors to multiply and expand geographically vector-borne diseases like dengue, zika, and others[72].

All the factors mentioned above, associated with conflicts, can collapse the control of NTDs due to the disruption of the health system, with disastrous consequences[73].

B.3 Data and Infrastructure Gaps

Weak surveillance systems, fragmented data streams, and inadequate health infrastructure in many endemic regions substantially limit decision-making, resource allocation, and the monitoring of control efforts. Solitary or narrowly focused interventions implemented in areas with poor infrastructure and fragmented data systems often lead to inefficient use of resources and failure to detect recrudescence or residual transmission[60,61]. Country-level reviews have documented examples where insufficient funding, capacity, or intensity of interventions contributed to resurgence or delayed elimination (reported in several settings including Equatorial Guinea, Egypt, Mexico, India), or where effective interventions were not fully implemented (for example Maldives, Ecuador, Guatemala, Bangladesh)[61].

Global estimates produced by the WHO (e.g., numbers of people requiring preventive chemotherapy or treatment) provide useful benchmarks but remain theoretical in many high-transmission or adult-predominant settings where target coverage thresholds derived from those estimates are insufficient to interrupt transmission[62,74]. Heterogeneous age–infection profiles and frequent co-infections further complicate programmatic decisions and may require tailored approaches rather than one-size-fits-all mass drug administration strategies[53]. These limitations make intensified post-treatment surveillance, integrated data systems, and investments in routine surveillance capacity essential components of any credible elimination agenda; without such investments, confirmation of elimination and the detection of residual infection will remain unreliable[60,64].

Supplementary Information C.

We searched Scopus and Web of Science for peer-reviewed research articles published between Jan 1, 2015 and Feb 20, 2025.

C.1 Scopus

```
TITLE-ABS-KEY ( ( "machine learning" OR "deep learning" OR "artificial intelligence"
OR "AI" OR "ML" OR "neural networks" OR "deep neural networks" OR "artificial
neural networks" OR "supervised learning" OR "unsupervised learning" OR "reinforcement
learning" OR "federated learning" OR "self-supervised learning" OR "foundation
models" OR "causal representation learning" OR "multimodal learning" ) AND (
"neglected tropical diseases" OR "NTDs" OR "malaria" OR "schistosomiasis" OR
"leprosy" OR "leishmaniasis" OR "lymphatic filariasis" OR "dengue fever" OR "Chagas
disease" OR "soil-transmitted helminths" OR "trypanosomiasis" OR "onchocerciasis"
OR "trachoma" OR "rabies" OR "buruli ulcer" OR "yaws" OR "chikungunya" OR "zika"
OR "snakebite envenoming" ) AND ( "diagnosis" OR "prognosis" OR "treatment monitoring"
OR "drug discovery" OR "disease surveillance" OR "outbreak prediction" OR "policy
planning" OR "healthcare" OR "public health" OR "epidemiology" OR "clinical decision
support" OR "personalized medicine" OR "biomarkers" OR "genomics" OR "medical
imaging" OR "wearable devices" OR "electronic health records" OR "EHR" OR "telemedicine"
) )
```

C.2 Web of Science (WoS)

```
("machine learning" OR "deep learning" OR "artificial intelligence" OR "AI" OR
"ML" OR "neural networks" OR "deep neural networks" OR "artificial neural networks"
OR "supervised learning" OR "unsupervised learning" OR "reinforcement learning"
OR "federated learning" OR "self-supervised learning" OR "foundation models"
OR "causal representation learning" OR "multimodal learning") AND ("neglected
tropical diseases" OR "NTDs" OR "malaria" OR "schistosomiasis" OR "leprosy" OR
"leishmaniasis" OR "lymphatic filariasis" OR "dengue fever" OR "Chagas disease"
OR "soil-transmitted helminths" OR "trypanosomiasis" OR "onchocerciasis" OR "trachoma"
OR "rabies" OR "buruli ulcer" OR "yaws" OR "chikungunya" OR "zika" OR "snakebite
envenoming") AND ("diagnosis" OR "prognosis" OR "treatment monitoring" OR "drug
discovery" OR "disease surveillance" OR "outbreak prediction" OR "policy planning"
OR "healthcare" OR "public health" OR "epidemiology" OR "clinical decision support"
OR "personalized medicine" OR "biomarkers" OR "genomics" OR "medical imaging"
OR "wearable devices" OR "electronic health records" OR "EHR" OR "telemedicine")
```

C.3 PRISMA 2020 Main Checklist

Topic	No.	Item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	Title; Abstract; Section 1, Page 2, Last Line
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	Section 1, Line 3, Line 6

continued on next page

Topic	No.	Item	Location where item is reported
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	Section 1, Page 3
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	Section 4.2, Line 1
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	Section 4.1
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	Section 4.1
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	Section 4.3; Section 4.4
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	Section 4.4, Line 9–14
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	Section 4.4, Line 15–18
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	Section 4.4, Line 19–22
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	Section 4.4, Line 15–22
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	Section 4.4, Line 15–18
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item 5)).	*NA*
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	*NA*

continued on next page

Topic	No.	Item	Location where item is reported
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	Section 2.1, Section 2.4
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	*NA*
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	*NA*
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	*NA*
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	*NA*
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	*NA*
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	Section 4.3
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	Section 4.3
Study characteristics	17	Cite each included study and present its characteristics.	Section 4.3
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	*NA*
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	Section 2.4
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	*NA*
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	Section 2.4
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	*NA*
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	*NA*
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	*NA*

continued on next page

Topic	No.	Item	Location where item is reported
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	*NA*
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	Section 2.4, Section 4.2
	23b	Discuss any limitations of the evidence included in the review.	Section 3, Line 16–31
	23c	Discuss any limitations of the review processes used.	Section 3, Page 15, Line 1–11
	23d	Discuss implications of the results for practice, policy, and future research.	Section 3, Page 14, Line 5–33
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	*NA*
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	*NA*
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	*NA*
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	Page 17
Competing interests	26	Declare any competing interests of review authors.	Page 17
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	Page 17

C.4 PRISMA 2020 Abstract Checklist

Topic	No.	Item	Reported?
TITLE			
Title	1	Identify the report as a systematic review.	Yes
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses.	Yes
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review.	Yes
Information sources	4	Specify the information sources (e.g. databases, registers) used to identify studies and the date when each was last searched.	Yes
Risk of bias	5	Specify the methods used to assess risk of bias in the included studies.	Yes
Synthesis of results	6	Specify the methods used to present and synthesize results.	Yes
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarize relevant characteristics of studies.	Yes
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each. If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e. which group is favored).	Yes
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g. study risk of bias, inconsistency and imprecision).	Yes
Interpretation	10	Provide a general interpretation of the results and important implications.	Yes
OTHER			
Funding	11	Specify the primary source of funding for the review.	Yes
Registration	12	Provide the register name and registration number.	Yes

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al[75]. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. MetaArXiv. 2020, September 14. DOI: 10.31222/osf.io/v7gm2. For more information, visit: http://www.prisma-statement.org

Supplementary Information D.

Table S4. Comparative Analysis of AI/ML Models for NTDs Diagnosis

Authors (Group)	NTDs	AI/ML Model(s)	Data Type	Region	Performance Range [%]	Key Limitations
Clinical ML[76–78]	Dengue	SVM, RF, XGBoost, ANN	Clinical, hematology datasets	Pakistan, Taiwan	Acc 72–100; Sens 80–100; Prec 15–70	Small datasets; class imbalance; limited external validation
Microscopy CNN[79–81]	Dengue	CNN, MobileNet, YOLOv8, SVM	Blood-smear microscopy	India, Australia	Acc 95–99.3; Sens 97–100	Single-site microscopy; device-specific variability
Microscopy + Vector ID[82–84]	Chagas	ResNet variants, U-Net, Boosting, CNN	Blood smears, histology, vector images	Mexico, Brazil, USA	Acc 98–100; Prec up to 100	Murine models; uneven class distribution; limited generalizability
Microscopy DL[85–87]	Leishmaniasis	YOLOv5, LeishFu-Net, Deep CNNs	Giemsa slides, bone marrow microscopy	Ethiopia, Iran, Brazil	Acc 95–99; Sens up to 100; Spec up to 98	Low sample size; monocyte-rich slides reduce performance
Microscopy + Digital Platforms[88,89]	Schistosomiasis	Faster R-CNN, AI digital microscopy	Urine/fecal microscopy; Schistoscope imaging	Nigeria, Gabon, Brazil	Sens 80–96; Spec 86–95; mAP up to 76.5	Underestimation at high egg load; low-contrast images
Imaging, Radiomics and Omics[90–92]	Schistosomiasis	Radiomics (SVM, RF, XGB), Omics regression, Spectral ML	Ultrasound radiomics, humoral, antibody profiles, MALDI-TOF spectra	China, USA, Brazil, Kenya, Europe	AUC 0.89–0.99; Acc 100 reported	Limited multi-center datasets; instrument variability
Microscopy + Mobile DL[93–97]	STH, Trachoma, Leprosy	Faster R-CNN, AutoML CNNs, L-CBAM, RF, Elastic-Net	Fecal smears, smartphone eyelid images, skin lesions	Philippines, Niger/Ethiopia, Brazil	Acc 88–97; Sens 84–95; Spec 87–95	Device variability; limited multi-country training
Diagnostic Support[98–101]	Rabies, Schistosomiasis-support	XGBoost, LR, NNAR, RF, ENVINet5	Surveillance data, remote sensing, ecological images	Haiti, Thailand, China	Sens 90; Spec 94; AUC up to 99.3	Lack of individual-level diagnostic data; ecological scaling limits

Supplementary Information E.

Table S5. Comparative Analysis of AI/ML Models for NTDs Prognosis

Study	NTDs	AI/ML Model(s)	Data type	Region	Performance Range	Key Limitations
[102]	Dengue	CART, Random Forest, Logistic Regression	Clinical + biochemical biomarkers	Bangladesh	Predictive probability 7–92.9	Single-centre; limited generalizability; cross-sectional
[103]	Dengue	Logistic Regression	Early hematology (platelets, WBC, MPV)	Northern India	Sensitivity = 90; Specificity = 70	Retrospective; missing data; local cohort only
[104]	Schistosomiasis (advanced)	LR, ANN, Decision Tree	Clinical exam + surveillance data	Hubei, China	AUC 0.82–0.93 (train); 0.81–0.83 (validation)	Small sample; model drift across sites
[105]	Cutaneous leishmaniasis	MLP, KNN, SVM, LVQ	Clinical lesion descriptors	Iran (Kerman)	Acc 87.8; Sens 90.3; Spec 86; AUC 0.88	Feature sparsity; possible overfitting; regional cohort
[106]	Snakebite envenoming	Bayesian Belief Networks	Clinical variables (pre/post antivenom)	Multi-country / SC2i	High interpretability; strong causal clarity	Limited to hospital-based cohorts; heterogeneity in reporting
[91]	Schistosomiasis	Interpretable ML + antibody profiling	Deep humoral profiling (immunology)	Africa/Asia (multi-site samples)	High class-separation; robust feature interpretability	Need longitudinal validation; immunoassay heterogeneity
[107]	Schistosomiasis	MALDI-TOF + ML classifiers	Mass spectrometry profiles	China/Africa reference libraries	Acc >90 (species differentiation)	Primarily diagnostic; limited prognostic linkage
[90]	Schistosomiasis fibrosis	Radiomics + ML	Ultrasound radiomics	China	AUC 0.80–0.90	Pilot sample size; imaging-standardization issues
[88]	Schistosomiasis (haematobium)	CNN / digital microscopy	Microscopy images (urine slides)	Gabon	High sensitivity/specificity (egg detection)	Diagnostic focus; prognostic use requires follow-up
[108]	Dengue	LSTM / sequence models	Time-series surveillance data	China (20 cities)	Reduced MAE/RMSE vs classical models	Sensitive to reporting lags; requires dense surveillance
[109]	Dengue	Hybrid Digital Twin + ML	Operational + environmental data streams	Case studies (Asia/Latin America)	Actionable lead time (weeks)	Pilot implementation; high engineering complexity

Table S5 – continued

Study	NTDs	AI/ML Model(s)	Data type	Region	Performance Range [%]	Key Limitations
[110]	Dengue	RBF networks + Darts Game Optimizer	Time-series + climate variables	India/Sri Lanka regions	Acc >85; improved stability metrics	Hyperparameter sensitivity; limited validation
[105]	Cutaneous leishmaniasis	ANN-based relapse prediction	Clinical data + treatment response	Pakistan + Iran	Acc range 80–92	Small training cohorts; class imbalance
[111]	Leprosy	SVM, Random Forest, XGBoost, CNN, LSTM	Genomic, clinical, imaging	India, Brazil, Pakistan	80–95 (various metrics)	Small datasets; class imbalance; regional specificity
[94]	Trachoma	AutoML CNN (on-device)	Smartphone images	Ethiopia / Saudi Arabia	Acc 85–92	Edge-device variability; limited longitudinal outcomes
[84]	Chagas (vector-based prognosis)	SVM, RF, ANN, CNN	Vector images (smartphone)	Latin America	Acc 70–95	Orientation sensitivity; model robustness issues
[80]	Dengue	Wavelet scattering + YOLO	Peripheral blood microscopy images	Multi-center experimental	Strong morphological class-separation reported	Very early-stage; limited sample size
[100]	Schistosomiasis (environmental risk)	Random Forest	Environmental + textural indicators	China (marshland/lake regions)	Accuracy >85	Environmental-only model; indirect prognostic utility

Supplementary Information F.

Table S6. Comparative Analysis of AI/ML Models for NTDs Treatment Monitoring

Authors	NTDs	AI/ML Model(s)	Data type	Region	Performance/Metric	Key Limitations
[50]	Leishmaniasis	DLNMs, XGBoost	Vector density, climatic data	Sri Lanka	Relative Risk: Rainfall(3.76), relative humidity(2.14),ambient temperature(2.81)	Limited exposure range, sampling bias, constrained parameter estimates
[51]	Cutaneous Leishmaniasis	Support Vector Machines	Clinical, sociodemographic	Iran	Accuracy 86.3, False-Negative Rate (FNR): 4% (Responsive) and 42% (Unresponsive)	Geographic limitation, random selection of sample size, lack of comparison
[70]	Dengue	BN, SVM, RBF tree, decision table, naive Bayes	Climate, meteorological data	Malaysia	Accuracy 92.35	Limited climatic factors considered, missing data
[71]	Dengue	PCA, K-Nearest Neighbor	Spectral biochemical data	Malaysia	Accuracy, Precision, Sensitivity, and Specificity (100%)	Lab-only samples, lack of real-world validation

Table S6 – continued

Authors	NTDs	AI/ML Model(s)	Data type	Region	Performance/Metrics	Key Limitations
[72]	Schistosomiasis	random forest	household and agricultural surveys	Burkina Faso	Overal GDP loss of 0.8%	Study focused exclusively on Burkina Faso, limited consideration of factors
[73]	Onchocerciasis	CART, logistic regression, BIC	Clinical and parasitological data	Cameroon	Odd ratio (46.7; 95% CI, 20.1–116.8; P < .001), false positive (1%), false negative (27.5%)	False positive based on the Ov16 RDT being negative, skin snips may not be achieved especially on large sample

Supplementary Information G.

Table S7. Comparative Analysis of AI/ML Models for NTDs Drug Discovery

Authors	NTDs	AI/ML Model(s)	Data type	Region	Performance Range [%] [metric]	Key Limitations
[112]	Onchocerciasis	Luminex-based antigen detection + statistical biomarker analysis (no AI/ML models)	Host serum biomarkers (IgG responses, peptide arrays, antigen signatures)	Ghana, Cameroon, DRC, Sudan, Uganda	Not applicable (diagnostic biomarker discovery; no ML metrics reported)	Not a drug-discovery or ML modelling study; biomarkers require validation across wider endemic regions; potential cross-reactivity with other filarial infections.
[113]	Schistosomiasis	YOLOv5 (object detection)	Brightfield microscopy images of *S. mansoni* schistosomula (4390 images; healthy vs. damaged)	Spain	mAP = 96.6%; Precision = 94.6%; Recall = 91.7% (internal validation)	Small-to-medium dataset; no motility assessment (phenotype only); external validation limited to image-based counting; no benchmark comparison with other DL models; requires further evaluation on diverse imaging setups.
[114]	Schistosomiasis	GNN-DDAS (GNN + MLP feature fusion with GNNExplainer)	Molecular graphs + SMILES-derived sequence features	China (Dali/Yunnan)	Not reported (abstract states superior to benchmarks)	Numerical metrics not provided in abstract; performance dependent on training data diversity; requires prospective wet-lab validation.

Table S7 – continued

Authors	NTDs	AI/ML Model(s)	Data type	Region	Performance Range [%] [metric]	Key Limitations
[115]	Leishmaniasis (*L. infantum*)	Multitask Deep Learning (MT-DNN, MT-MPNN)	Cheminformatics descriptors (ECFP4), molecular graphs, ChEMBL bioactivity data; virtual screening of ∼1.3 M ChemBridge compounds	Brazil, France, USA	Classification (random split): ACC = 0.80 ± 0.02 for *L. infantum*; Regression: r = 0.83 ± 0.06, RMSE = 0.45 ± 0.05, pIC_{50}	Model performance dependent on data imbalance; scaffold split performance markedly lower; limited chemical diversity for *L. infantum*; requires further toxicity profiling and hit-to-lead optimization.
[116]	Schistosomiasis	2D/3D QSAR (HQSAR, CoMFA, CoMSIA; consensus modeling)	Cheminformatics descriptors + curated IC_{50} dataset + virtual screening outputs	Brazil/USA/UK collaboration	RMSEP = 0.12˜0.18, Q^2_{ext} = 0.90˜0.98	Model depends on correct molecular alignment; scaffold diversity limited in training set; 3D-QSAR sensitive to charge assignment; prospective validation needed for generalization.

Supplementary Information H.

Table S8. Comparative Analysis of AI/ML Models for NTDs Surveillance

Authors	NTDs	AI/ML Model(s)	Data Type	Region	Performance Metrics	Key Limitations
[117]	Dengue	CIF, RF, XGB, GAM, ANN, SVM (Hybrid)	Remote sensing, case data	Philippines	HAI=0.771; AUC=0.794; S-R^2=0.716	Zero-inflated data (>70%); limited spatial variance; Metro Manila only
[118]	Dengue	RF, SVR, DT, MLR, ARIMA, SARIMA, Prophet	Climate (temp, humidity, rainfall)	India (9 cities)	R^2=0.84 (Prophet, Mumbai)	Only climatic factors; monthly granularity; 9 cities only
[119]	Dengue	ANN (5 layers), ARIMA(6,1,1)	Monthly case data	Bangladesh (Dhaka)	ANN: RMSE = 7.59×10^6, MAPE=1.15%; ARIMA: RMSE=0.0311	Single city; no climate predictors; overfitting risk
[120]	Dengue	CatBoost, LSTM, SVM, RF (Ensemble)	Climate, case data	Brazil, Peru	NRMSE varies; 1-month forecasts	Difficulty with extreme values; data availability dependency

Table S8 – continued

Authors	NTDs	AI/ML Model(s)	Data Type	Region	Performance Metrics	Key Limitations
[121]	Dengue	CNN, Transformer, LSTM, LSTM-ATT	Meteorological, case data	Vietnam (20 provinces)	LSTM-ATT: RMSE=1.60, MAE=1.95	Excludes non-weather factors; missing data
[122]	Dengue	XEWNet (ensemble wavelet NN)	Climate, case data	Puerto Rico, Peru, India	52wk: RMSE=9.98, MAE=6.55, SMAPE=78.34	Limited dataset size; geographic diversity constrained
[123]	Dengue	NBR, SARIMAX, XGBoost, LSTM	Meteorological (2003-2022)	Vietnam (BRVT)	NBR: MAE(test)=20.31, RMSE(test)=27.19	Limited short-term peaks; convergence issues
[124]	Dengue	GAMLSS, Random Forest	Climate, VAR predictions	Costa Rica (32 cantons)	NRMSE, NIS (not specified)	Limited capacity in some cantons
[110]	Dengue	RBFN-DGO	Climate, case data	India, Peru, USA	Acc=99.98%; MAE=25.98%; RMSE=29.12%	Data inconsistency; limited generalization
[125]	Dengue	RF, DT, GNB, SVC, KNN (DDPM)	Sentiment, epi data	Puerto Rico, Peru	Mean scores: RF=8.72, KNN=6.31	High computational demand; poor data quality
[126]	Dengue	Hybrid GANN (GA+NN)	Case data, neighboring locations	Malaysia (Selangor)	MSE=0.0199-0.0384; R^2=0.8761-0.9382	Small dataset; excludes environmental factors
[127]	Dengue	Bayes-PSO	Climate (TempWind)	India	Acc=97.5%	Limited dataset; uncertain generalizability
[128]	Schistosomiasis	LR, Elastic Net, RF, XGBoost, SVM, MLP	Epi surveys, remote sensing	Sub-Saharan Africa (5 countries)	Acc=90%; Sens=90%; Spec=90%	Limited generalizability; uneven samples
[129]	Schistosomiasis	CNN-IDE, IDE, ST Kriging	Parasitological, environmental	China (Anhui)	MSPE=0.04; CRPS=0.19	Diagnostic limitations; limited covariates
[130]	Rabies	Imbalanced-XGBoost, LR, SVM, NN, RF	Geospatial kriging, socio-economic	Morocco (1546 communes)	Acc=92%; F1=92%; 25.8% high-risk	Underreported data; missing covariates
[111]	Leprosy	Bayesian Networks	Clinical, demographic, genetic	Brazil (1450 patients)	Acc=87.7%; Sens=85.7%; Spec=89.4%	Limited genetic data access
[131]	Scabies	Feedforward NN, Stochastic ODE	Synthetic, simulated	N/A (theoretical)	R_0 stability analysis	No real-world validation; computational cost

Supplementary Information I.

Table S9. Comparative Analysis of AI/ML Models for NTDs (Grouped by Model Family, with References)

Authors	AI Model(s)	Data Types	Regions	Performance Metrics & Key Limitations
[108, 132–137]	LSTM, RNN, GRU	Climate; Demographic; Epidemiological; Meteorological; Environmental; Spatiotemporal	Malaysia; India; Bangladesh; Indonesia; China	Accuracy: 70–97%; RMSE: 1–693; AUROC: 0.95; F1 score: 57–97%. Limitations: small datasets, overfitting, black-box nature, high computational cost, missing influential factors, limited generalizability.
[138, 139]	ANN, Perceptron, Hybrid NN	Environmental; Demographic; Epidemiological; Pathogen Serotype	USA (Puerto Rico); Mexico (Yucatan); Bangladesh	Accuracy: 70–97%; convergence issues; overfitting. Limitations: missing influential factors, dependence between geographical and demographic factors.
[140–144]	Tree-based Models (DT, RF, XGBoost, Gradient Boosting)	Epidemiological; Climate; Geographical; Meteorological	Bangladesh; Brazil; China; Vietnam	Accuracy: 74–97%; AUC: 0.76–0.94; MAE / RMSE: 32–83. Limitations: overfitting, assumption of independence, complex interpretation, missing key features.
[142, 145]	SVM, KNN, Logistic Regression, Naïve Bayes, LogitBoost	Epidemiological (clinical features); Climate; Meteorological	India; Bangladesh	Accuracy: 90–92%; Sensitivity: 90%; Specificity: 94%; F1 Score: 92%. Limitations: small datasets, possible overfitting, high computational cost, less interpretability.
[137, 143, 146]	PT-Ensemble, SARIMA, Prophet, Hybrid Models	Climate; Epidemiological; Meteorological; Environmental	Bangladesh; Vietnam	Accuracy: 80–89%; MSE: 2.61–8.59; RMSE: 1.61–1.089; MAE: 610.09. Limitations: data fragmentation, oversimplified disease dynamics, high computational cost.
[147, 148]	GWR, Fuzzy-Logistic Regression	Demographic; Geographical; Epidemiological; Climate	Indonesia	MSE: 8.59; Accuracy: 79–79.93%. Limitations: small datasets, collinear features, outliers, subjective fuzzification, limited generalizability.

Continued on next page

Table S9 – continued

Authors	Model Family	Data Types	Regions	Performance Metrics & Key Limitations
[140, 149]	RF, SDM (Schistosomiasis)	Climate; Geographical; Environmental	Brazil; China	AUC: 0.76–0.94; Kappa: 0.618. Limitations: assumes independence, ignores interactions, missing critical factors (snail species, human behavior), old datasets.

Supplementary Information J.

Table S10. Comparative Analysis of AI/ML Models for NTDs Policy Planning

Authors	NTDs	AI/ML Model(s)	Data type	Region	Performance Range	Key Limitations
[150]	Dengue	Not specified	N/A	N/A	N/A	N/A
[151]	Dengue	Not specified (simulation-based)	N/A	N/A	N/A	N/A
[152]	Leishmaniasis	Not specified (perspective)	Perspective (policy)	Canada	No specific evaluation metrics	Challenges in partnering and licensing
[153]	Schistosomiasis	XGBoost, SVM, GLM, Regression Tree, Random Forest, GBM, Neural Network	Epidemiological, cost, economic databases	China	R^2=0.7306, MSE=0.0041, RMSE=0.0641, MAE=0.0053	Limited to four provinces; limited generalizability
[154]	Trachoma	PA, Cohen's κ	Photographs	Ethiopia	PA = 96.7%; κ = 0.70 (95% CI 0.64–0.77)	SLR camera not scalable
[155]	Schistosomiasis	R statistical analysis	Cross-sectional survey (humans)	Nigeria	Sensitivity 67.7%; Specificity 100%	Limited metric reporting
[156]	Dengue	NB, DT, RF, SVM, XGB, MLP, CNN, LSTM, BiLSTM, DuBiLSTM	ProMED-mail (volunteer reporting)	Southeast Asia	F1=93.34%; Precision=92.42%; Recall=94.38%	Inconsistent reporting; difficulty identifying dengue in multiclausal text
[157]	Dengue	Classification models	Literature review	N/A	Qualitative assessment	No detailed evaluation metrics
[158]	Dengue	Deep learning models	Health system data	N/A	Classification accuracy (not explicitly stated)	Not explicitly stated
[159]	Lymphatic filariasis	RF, QRF	Prevalence surveys	Nigeria	R^2=0.4, RMSE=1.24	Poor performance for extreme quantiles in small datasets

Table S10 – continued

Authors	NTDs	AI/ML Model(s)	Data type	Region	Performance Range	Key Limitations
[160]	Leishmaniasis	None (composite indicator)	SisLeish (online surveillance)	Latin America (16 countries)	SWOT analysis (no metrics)	Underreporting, missing data, outdated maps

Supplementary Information K.

K.1 Datasets and Data Repositories for AI-Based NTD Research

The rapid expansion of AI applications for NTDs has heightened the importance of dataset availability, quality, and transparency. Across systematic reviews, existing datasets are consistently reported to be small, imbalanced, poorly documented, and rarely publicly accessible, limiting reproducibility and generalizability[161–163]. Many studies omit critical information on data provenance, preprocessing, and model configuration, and more than half do not release datasets, reinforcing structural barriers to independent validation. Moreover, AI performance is constrained by heavy reliance on non-patient or secondary data, with relatively few studies using patient-level clinical records, despite their importance for clinically grounded modeling[162].

Despite coordination gaps, government agencies and international institutions remain central producers of NTD-relevant data. Surveillance systems, registries, hospitals, laboratories, and government meteorological services underpin most dengue forecasting and prediction models[164,165], challenging claims of limited governmental involvement[166]. Registry-based and population-level datasets dominate (83.8%), with temporal resolutions ranging from monthly to daily, highlighting the importance of high-frequency data for capturing outbreak dynamics[165]. Persistent challenges include underreporting, misclassification, and inconsistencies between global and national records, as demonstrated for dengue surveillance data[164].

A diverse ecosystem of global, disease-specific, and general-purpose repositories supports AI-driven NTD research. Major platforms include the WHO Global Health Observatory, Global Health Data Exchange, ECDC surveillance systems, the ESPEN portal for African NTDs, and disease-focused initiatives such as GARC and historical atlases like GAHI. Dengue research benefits from particularly mature resources spanning epidemiology, molecular biology, and immunology, including DengueNet, the Dengue Outbreak Database[167], DENVDB, the Dengue Drug Target Database, DENV-Ab[168], and ViPR[169]. Complementary environmental and population-level data are sourced from open government portals, Kaggle, the UCI Machine Learning Repository, national health platforms, and the WMO, though heterogeneity in formats and metadata remains a barrier[170]. Integrated epidemiological resources such as the FAO/PAAT trypanosomiasis atlas illustrate the value of centralized, standardized, long-term data infrastructures for scalable and reproducible AI applications[171].

Table S11. Representative datasets supporting AI/ML research in neglected tropical diseases (NTDs), grouped by thematic categories.

Dataset Portal	Disease Area	Data Type	Access	Strengths	Limitations
Global Health & Burden-Estimation Portals					
WHO GHO / GHDx	Multiple NTDs	Surveillance indicators, burden estimates, DALYs	GHO / GHDx	Standardized global metrics; widely validated	Limited granularity; modeled estimates
WHO DengueNet	Dengue	Case counts, CFR, serotypes	WHO	Standardized global reporting	Underreporting in several regions

Continued on next page

Dataset Portal	Disease Area	Data Type	Access	Strengths	Limitations
GIDEON	Multiple infectious diseases	Epidemiological time series, outbreaks	GIDEON	Validated global epidemiology	Subscription-based; export restrictions
Regional / National Surveillance Platforms					
ESPEN (WHO-AFRO)	Schistosomiasis, LF, Onchocerciasis	Subnational surveys, prevalence maps	ESPEN	High-resolution African datasets	Africa-only coverage
ECDC Surveillance Database	Dengue, emerging NTDs	Case counts, outbreak intelligence	ECDC	Structured European surveillance	Limited relevance outside Europe
OGD India / Data.gov	VBDs, climate-linked NTDs	Surveillance, climate, demographics	OGD India / Data.gov	Large-scale datasets; open access	Heterogeneous formats; requires cleaning
Disease-Specific Epidemiological Atlases					
FAO/PAAT Trypanosomiasis Atlas	Trypanosomiasis	Entomological and animal infection data	FAO	Standardized Africa-wide atlas	Region-specific
GARC	Rabies	Surveillance, case investigations	GARC	High-quality rabies data	Geographic gaps; incomplete reporting
GAHI	Helminth infections	Spatial prevalence maps	GAHI	Model-ready geospatial layers	Last major update in 2015
Dengue-Focused Specialized Databases					
DENVDB	Dengue	Genomic sequences, outbreak metadata	DENVDB	Rich genomic and serotype information	Limited clinical metadata
Dengue Outbreak Database	Dengue	Historical outbreak records	DOI	Comprehensive historical dataset	No updates after 2012
DENV-Ab Database	Dengue	Antibody epitope interactions	DENV-Ab DB	Curated immunological data	Highly specialized
Machine-Learning-Ready Repositories					
Kaggle VBD Datasets	Dengue, Zika, Chikungunya	ML-ready structured datasets	Kaggle	Easy access; ML-friendly	Variable curation quality
UCI ML Repository	VBD-related datasets	Structured tabular data	UCI	Benchmark ML datasets	Limited NTD-specific resources
Climate & Environmental Data Sources					

Continued on next page

Dataset Portal	Disease Area	Data Type	Access	Strengths	Limitations
WMO Climate Data	Climate-linked NTDs	Temperature, rainfall, humidity	WMO	High-quality meteorological data	Not disease-specific

K.2 Tools and Frameworks

Effective deployment of AI for NTDs relies on a combination of machine learning and deep learning toolkits, digital health platforms, and explainable AI methods that support data collection, model development, and large-scale integration within health systems. These tools form the technical backbone of AI-driven solutions, enabling surveillance, diagnosis, prediction, and decision support, particularly in resource-constrained environments.

Modern AI applications frequently employ established machine learning and deep learning frameworks that provide modular, scalable, and well-documented environments for model development. TensorFlow[172] and PyTorch[173] are two of the most prominent deep learning libraries, offering extensive support for computer vision, natural language processing, and time-series analysis.

Classical machine learning remains essential for many NTD datasets. The scikit-learn library[174] provides efficient implementations of supervised and unsupervised learning algorithms suitable for risk prediction and epidemiological modeling.

AI solutions for NTDs depend heavily on high-quality health data. Digital health platforms such as District Health Information System 2 (DHIS2)[175] and OpenMRS[176] play a central role in organizing and standardizing data. Complementary tools include Open Data Kit (ODK)[177] for mobile data collection, and geospatial platforms such as Google Earth Engine[178] and QGIS[179].

Explainable AI (XAI) and fairness toolkits improve transparency and trust in model predictions. These methods help clinicians understand model behavior and ensure ethical deployment in health systems. Important frameworks include SHAP[180], LIME[181], Captum[182], TorchCAM[183], and iNNvestigate[184].

Clinical and fairness-oriented AI toolkits such as MONAI[185], Fairlearn[186], and AIF360[187] support trustworthy and clinically robust model development.

Table S12. AI/ML tools, digital health platforms, geospatial systems, fairness libraries, and XAI frameworks relevant to NTDs.

Tool	Function	Use Case	Open	Health Specific	Strength for NTDs
1. Deep Learning Frameworks					
TensorFlow[172]	Deep learning framework	Imaging, diagnostic AI	Yes	No	Strong deployment ecosystem.
PyTorch[173]	DL library	Classification, NLP	Yes	No	Flexible, research-friendly.
scikit-learn[174]	Classical ML	Risk models, tabular datasets	Yes	No	Ideal for limited NTD datasets.
MONAI[185]	Medical imaging AI	Segmentation, detection	Yes	Yes	Optimized for clinical imaging tasks.
2. Digital Health Platforms					
DHIS2[175]	Health data system	Surveillance, reporting	Yes	Yes	Standard in LMIC health systems.
OpenMRS[176]	EMR	Clinical record management	Yes	Yes	Enables structured patient datasets.

Continued on next page

Tool	Function	Use Case	Open	Health Specific	Strength for NTDs
ODK[177]	Mobile data capture	Field/community surveys	Yes	Partial	Offline capability for remote NTD regions.
3. GIS and Environmental Mapping Tools					
Google Earth Engine[178]	Geospatial computing	Env./climate mapping	Free	No	Useful for vector ecology modeling.
QGIS	GIS software	Spatial analysis	Yes	No	Offline GIS for epidemiology.
4. Fairness and Bias Assessment Toolkits					
Fairlearn[186]	Fairness assessment	Bias analysis	Yes	No	Detects demographic disparities.
AIF360[187]	Fairness toolkit	Bias metrics + mitigation	Yes	No	Includes multiple debiasing algorithms.
5. Explainable AI (XAI) Frameworks					
SHAP[180]	Attribution	Local/global explanations	Yes	Partial	Interpretable risk and severity models.
LIME[181]	Local explanations	Case-level interpretation	Yes	Partial	Good for clinical justification.
Captum[182]	PyTorch XAI	Network attribution	Yes	No	Native integration with deep learning models.
TorchCAM[183]	CAM visualization	Image-based explanation	Yes	No	Effective for lesion/rash detection.
iNNvestigate[184]	DL interpretability	Saliency, LRP	Yes	No	Supports multiple XAI techniques.

Together, these tools create an integrated ecosystem for AI-enabled NTD control: data collection platforms support surveillance, ML and DL frameworks power model development, and XAI and fairness toolkits ensure transparency, trust, and equitable deployment.

Supplementary Information L.

While standard paradigms (i.e., supervised, unsupervised, and reinforcement learning) underpin most current AI applications in NTDs, emerging methodological paradigms have the potential to fundamentally reshape how evidence is generated, models are trained, and insights are translated into action. Below, we critically examine how these paradigms could impact NTD research, highlighting both their transformative potential and their limitations in low-resource, high-heterogeneity settings.

L.1 Federated Learning (FL)

Figure S1 illustrates the FL paradigm in which models are trained across decentralized clinical sites without sharing raw patient data[188–190]. FL directly addresses one of the most persistent bottlenecks in NTD research: fragmented, small-scale, and institutionally siloed datasets. By enabling collaborative model training without centralizing data, FL can increase sample diversity across geographies, transmission settings, and health systems, thereby improving model generalizability and reducing site-specific bias[191]. This is particularly relevant for NTDs, where data are often distributed across ministries of health, NGOs, and regional hospitals.

However, FL introduces non-trivial challenges. Data heterogeneity across sites (non-IID data), limited local computational infrastructure, and unstable connectivity can degrade convergence and performance. Moreover, privacy is not absolute: model updates may still leak sensitive information through gradient inversion or membership inference attacks. These limitations

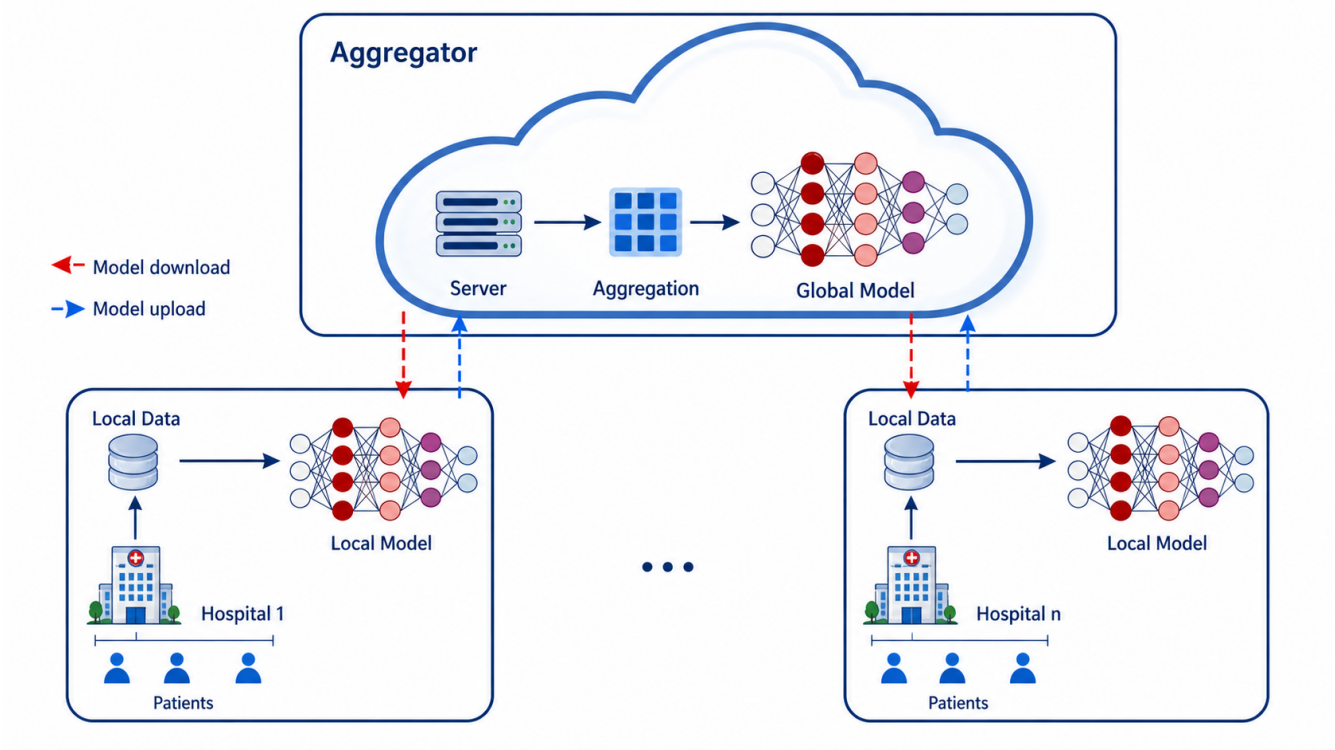


Figure S1. Overview of the FL framework. Local models are trained on data at individual hospitals or institutions (e.g., Hospital 1 to Hospital *n*) without sharing sensitive patient data. Model updates are aggregated on a central server to create a global model, ensuring privacy while enabling collaborative AI development across geographically distributed datasets. This approach is particularly beneficial for advancing healthcare applications in NTDs.

can be mitigated through personalized federated learning, robust aggregation strategies, communication-efficient training, and the integration of differential privacy and secure aggregation. Without such safeguards, FL risks reproducing inequities by privileging better-resourced sites.

L.2 Self-Supervised Learning (SSL)

Self-supervised learning has emerged as a core strategy for representation learning in domains characterized by limited annotation and high acquisition costs[192,193]. SSL offers a pragmatic pathway to leverage the vast amounts of unlabeled data that characterize NTD contexts, including microscopy images, ultrasound scans, and longitudinal surveillance data. By learning general-purpose representations before task-specific fine-tuning, SSL can substantially reduce annotation requirements and improve downstream performance in data-scarce settings[194–196].

Yet, SSL is not a panacea. Pretext tasks may encode spurious correlations or domain-specific artifacts that do not transfer across regions, pathogens, or imaging protocols. In NTDs, where data acquisition conditions could vary widely, SSL models risk learning shortcuts that undermine clinical validity. Mitigation strategies include domain-adaptive pretraining, careful validation across sites, and combining SSL with limited expert annotation to anchor representations in clinically meaningful features.

L.3 Foundation Models

Foundation models are increasingly viewed as general-purpose infrastructures for downstream biomedical AI applications[197,198]. Foundation models promise to amortize learning across tasks by capturing broad linguistic, visual, or molecular priors[197]. In NTD research, they could enable rapid development of diagnostic tools, literature synthesis, decision support, and drug discovery pipelines[199–205]. Their capacity to integrate heterogeneous knowledge sources is particularly attractive in fields where formalized datasets are sparse but tacit knowledge is abundant.

However, foundation models are predominantly trained on data from high-income settings, raising concerns about representational bias, domain mismatch, and unsafe extrapolation. Their computational cost and carbon footprint further limit accessibility in low-resource environments[206–209]. Meaningful impact in NTDs will therefore depend on targeted adaptation strategies, including domain-specific fine-tuning, parameter-efficient transfer learning, and governance mechanisms to ensure transparency, evaluation, and equitable deployment.

L.4 Causal Representation Learning

Causal representation learning (CRL) builds on a growing body of work at the intersection of causality and representation learning[210,211]. As shown conceptually in Figure S2, CRL seeks to recover latent causal structure from observational and interventional data. It addresses a core limitation of correlational AI systems: fragility under distribution shift. By explicitly modeling causal structure, these approaches can support more robust predictions and counterfactual reasoning in the presence of environmental change, intervention, or policy shifts[212]. For NTDs, this is particularly relevant for understanding transmission dynamics influenced by climate, vector ecology, and socio-economic factors.

Nonetheless, causal methods rely on strong assumptions and often require interventional or longitudinal data that may be scarce in NTD settings. Mis-specified causal graphs may yield misleading inferences, with direct implications for public health decision-making. Progress will require closer integration between causal modeling, domain expertise, and data collection strategies, including hybrid approaches that combine mechanistic models with data-driven learning.

L.5 Multimodal Learning

Multimodal learning has been widely adopted in biomedical AI to integrate heterogeneous clinical and biological data streams[213,214]. Figure S3 contrasts unimodal and multimodal learning, highlighting the benefits of integrating heterogeneous data sources in complex disease settings such as NTDs. Multimodal learning aligns naturally with the complexity of NTDs, where diagnosis and prognosis depend on the integration of clinical presentation, imaging, laboratory tests, and environmental context[215,216]. By jointly modeling these modalities, multimodal systems can improve robustness and reduce reliance on any single noisy data source.

However, multimodal approaches amplify challenges related to missing data, modality imbalance, and inconsistent data quality issues that are pervasive in NTD research. Models may overfit to dominant modalities or fail gracefully when key inputs are absent. Effective deployment therefore requires architectures that explicitly handle missingness, uncertainty-aware fusion strategies, and careful evaluation under realistic field conditions.

In synthesis, these emerging paradigms offer substantial opportunities to move AI in NTD research beyond narrow, task-specific solutions toward more generalizable, robust, and equitable systems. Realizing this potential will require not only technical innovation but also methodological rigor, context-aware evaluation, and sustained investment in data infrastructure tailored to the realities of NTD-endemic settings.

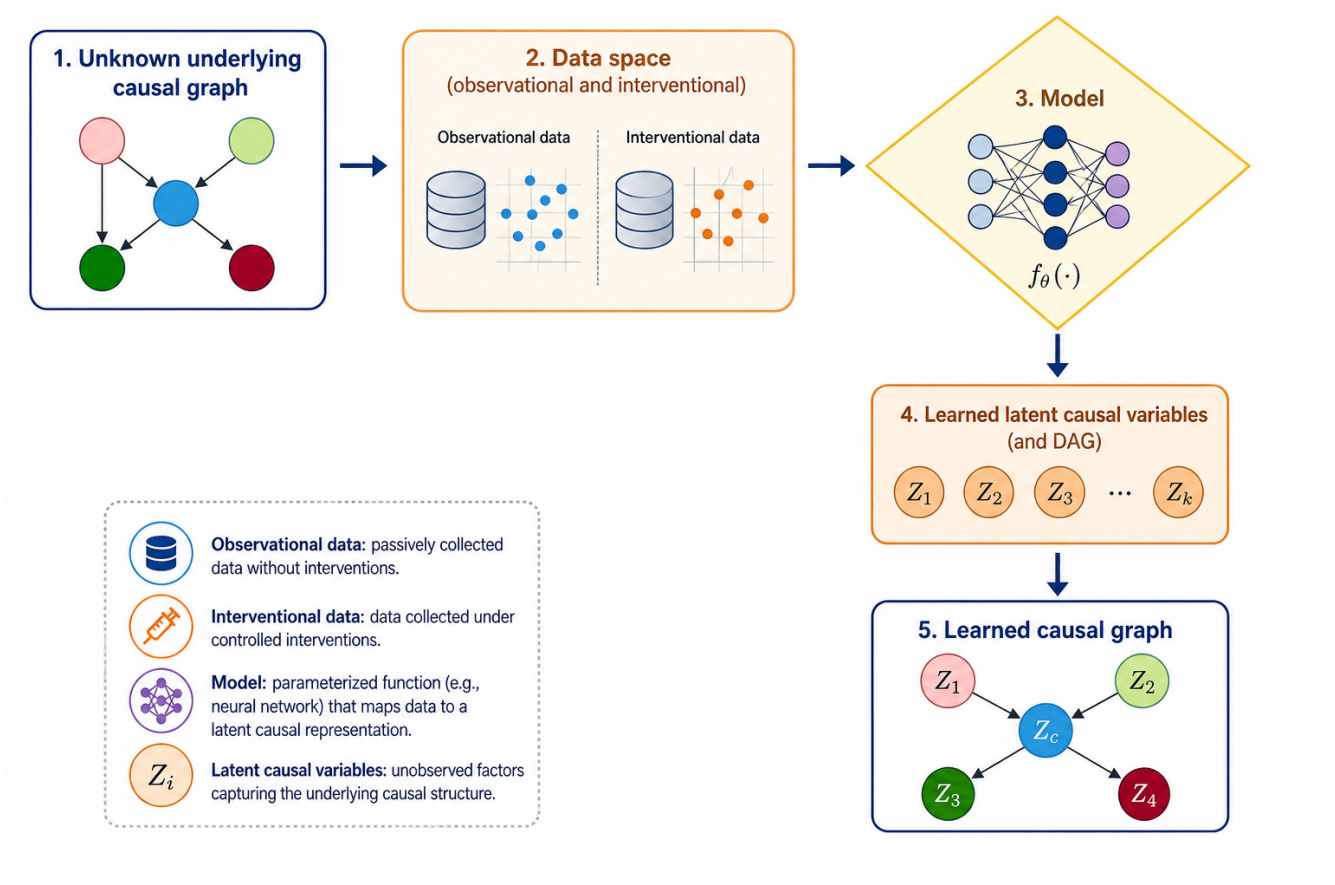


Figure S2. Framework for deep CRL in AI. The process begins with observational and interventional data, from which latent causal variables and their directed acyclic graph (DAG) are learned. This approach enables the discovery of robust causal relationships, improving prediction accuracy and adaptability in complex settings such as disease dynamics and treatment outcomes for NTDs.

Supplementary Information M.

AI applications for NTDs face interlinked data, technical, and ethical constraints that limit robustness, equity, and real-world impact.

M.1 Data Scarcity, Bias, and Representativeness

Persistent data scarcity and geographic bias remain core limitations. Most datasets originate from a small number of countries (notably Brazil, India, China, and high-income research hubs), while many other highly endemic regions in sub-Saharan Africa and Southeast Asia remain underrepresented. This reflects long-standing disparities in surveillance infrastructure, laboratory capacity, and digitization, resulting in AI models trained on small, homogeneous datasets that fail to capture ecological, genetic, and clinical diversity[24,217,218]. Consequently, models often generalize poorly when deployed in high-burden, low-resource settings, reinforcing inequities rather than mitigating them[219–221]. These challenges are exacerbated by limited longitudinal data, sparse annotations, non-standardized metadata, and fragmented reporting, constraining the development of robust multimodal and population-level models[217,222].

Disease-specific data availability further skews innovation. NTDs with relatively mature imaging or surveillance pipelines (e.g., dengue, leishmaniasis, schistosomiasis) dominate the literature, whereas conditions such as Buruli ulcer, mycetoma, chromoblastomycosis, and dracunculiasis are largely absent from public datasets[223–226]. This risks aligning AI research with existing digital infrastructure rather than unmet public health needs. Addressing these gaps requires coordinated investment in locally generated, representative datasets, strengthened routine surveillance, and decentralized learning paradigms (e.g., federated or swarm learning) that preserve data sovereignty while improving external validity[188,227,228].

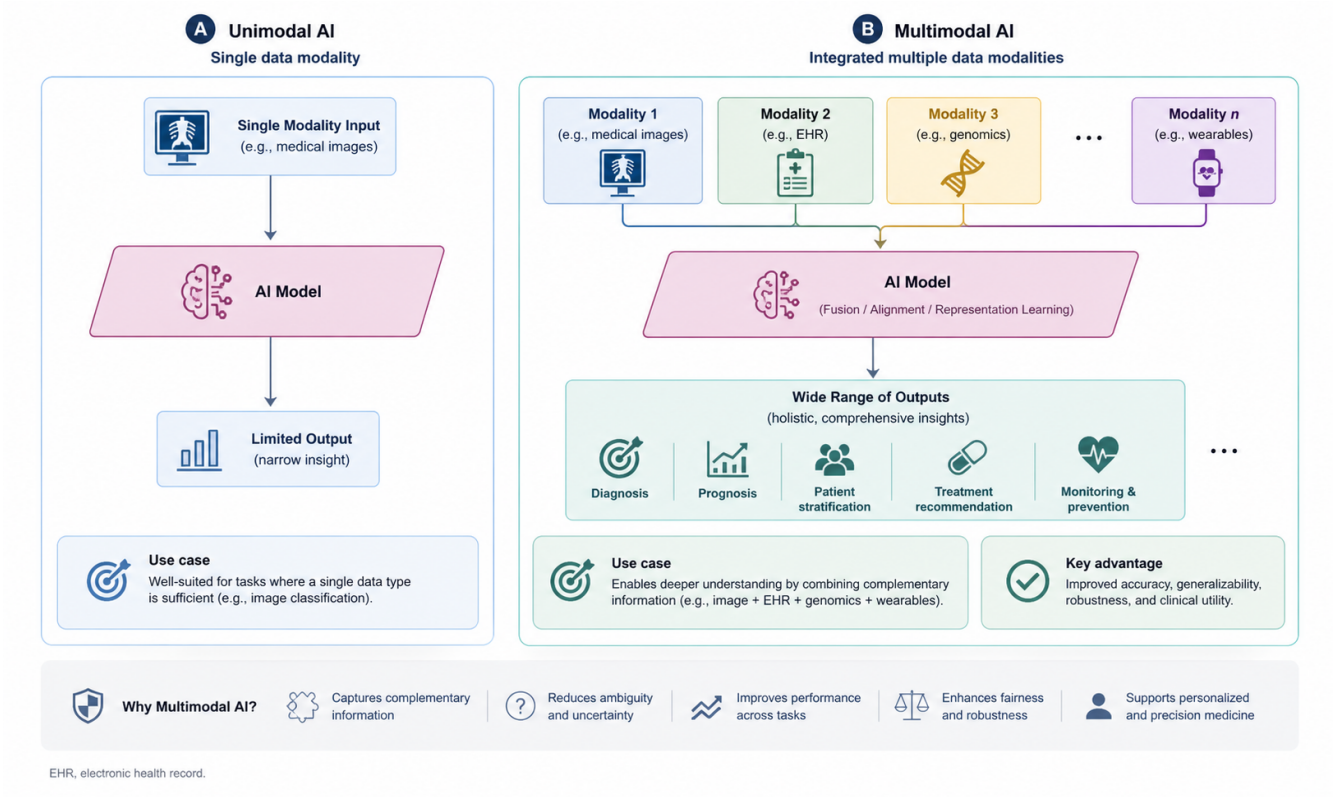


Figure S3. Comparison of unimodal and multimodal learning approaches in AI/ML. (a) Unimodal learning relies on a single data modality, such as images, providing limited output and insights. (b) Multimodal learning integrates multiple data modalities, such as images, audio, and text, enabling more comprehensive outputs and improved performance in complex tasks like diagnosis and prognosis in NTDs.

M.2 Technical and Infrastructural Barriers

Technical constraints remain significant in endemic, resource-limited settings. Many AI systems, particularly deep learning models, depend on high-performance computing, stable electricity, and reliable connectivity, which are often unavailable in low-resource or rural contexts[222,227,229–231]. These limitations affect not only model training but also deployment, updating, and long-term maintenance. Health-economic evaluations frequently overlook indirect costs such as infrastructure, maintenance, and workforce training, leading to overly optimistic cost-effectiveness claims[232,233].

Capacity gaps compound these issues. Many endemic-country health systems lack trained personnel in machine learning, data engineering, and IT support, increasing dependence on external expertise and undermining sustainability and local ownership[231,234]. Reproducibility and transparency are additional concerns: non-deterministic model behavior, sensitivity to preprocessing pipelines, and inconsistent reporting impede independent validation and trust[235–237]. Without rigorous external validation, detailed documentation, and open-science practices (where ethically permissible), transferability across settings remains uncertain.

M.3 Socioeconomic and Ethical Considerations

Beyond technical limitations, AI deployment for NTDs raises critical socioeconomic and ethical issues. Chronic underfunding of NTD programmes, weak digital infrastructure, and reliance on external donors constrain national capacity to develop, sustain, and scale AI systems, particularly in fragile or conflict-affected settings. AI-enabled diagnostics and surveillance rely on sensitive health data, raising concerns about privacy, cyber-security, and informed consent, especially in communities with limited data literacy. Even anonymized datasets may be re-identifiable, exposing already marginalized populations to stigma or

harm.

Unclear data ownership and governance further complicate implementation. Ambiguities regarding whether data belong to patients, local institutions, governments, or external developers risk enabling extractive practices and "surveillance capitalism," where public-health data are repurposed with limited local benefit. These dynamics intersect with broader concerns of digital colonialism: most AI tools are designed, hosted, and governed in high-income countries, positioning endemic regions as data providers rather than equal partners. Limited technology transfer, persistent digital divides, and global competition for AI expertise contribute to dependency, algorithmic bias, and brain drain, undermining equitable adoption and long-term impact.

M.4 Toward Responsible and Equitable AI for NTDs

Collectively, these challenges highlight the need for coordinated, equity-centered approaches to AI for NTDs. Federated data networks aligned with FAIR principles can enable privacy-preserving use of real-world data and generation of real-world evidence while improving external validity[238,239]. Existing initiatives, such as the European Union's Integration of Heterogeneous Data and Evidence towards Regulatory and HTA Acceptance (IDERHA), demonstrate the feasibility of federated, cross-border infrastructures with robust governance and interoperability standards[240]. Progress will require sustained investment in digital infrastructure, workforce capacity, transparent governance, and adherence to established AI reporting standards, ensuring that AI innovations support, rather than undermine, equitable NTD control.

References

1. Melo, G. *et al.* Evolution of research funding for neglected tropical diseases in brazil, 2004–2020. *PLoS Neglected Trop. Dis.* **17**, e0011134, DOI: 10.1371/journal.pntd.0011134 (2023).

2. WHO. R&d funding flows for neglected diseases by disease, year and funding category (2025). Accessed September 2025.

3. Júnior, J., Ramos, A., Alves, J. *et al.* Measuring social vulnerability in communities and its association with leprosy burden through spatial intelligence in central west brazil to guide strategic actions. *Arch. Public Heal.* **82**, 246, DOI: 10.1186/s13690-024-01484-1 (2024).

4. to Combat Neglected Tropical Diseases, U. Neglected tropical diseases directory (2025). Accessed September 2025.

5. Pellegrino, C., Patti, G., Camporeale, M. *et al.* Guinea worm disease: A neglected disease on the verge of eradication. *Trop. Medicine Infect. Dis.* **7**, 366, DOI: 10.3390/tropicalmed7110366 (2022).

6. Osei-Owusu, J., Aidoo, O., Eshun, F. *et al.* Buruli ulcer in africa: Geographical distribution, ecology, risk factors, diagnosis, and indigenous plant treatment options – a comprehensive review. *Heliyon* **9**, e22018, DOI: 10.1016/j.heliyon.2023.e22018 (2023).

7. Mahapatra, B., Mukherjee, N., Khatoon, S. *et al.* Economic evaluations of neglected tropical disease interventions in low- and middle-income countries: a systematic review protocol. *JBI Evid. Synth.* **22**, 1582–1593, DOI: 10.11124/JBIES-23-00339 (2024).

8. Hotez, P. & Kamath, A. Neglected tropical diseases in sub-saharan africa: Review of their prevalence, distribution, and disease burden. *PLoS Neglected Trop. Dis.* **3**, e412, DOI: 10.1371/journal.pntd.0000412 (2009).

9. Kirigia, J. & Kubai, P. Monetary value of disability-adjusted life years and potential productivity losses associated with neglected tropical diseases in the east african community. *Res. Reports Trop. Medicine* **14**, 35–47, DOI: 10.2147/RRTM.S382288 (2023).

10. Semahegn, A., Manyazewal, T., Getachew, E. *et al.* Burden of neglected tropical diseases and access to medicine and diagnostics in ethiopia: a scoping review. *Syst. Rev.* **12**, 140, DOI: 10.1186/s13643-023-02302-5 (2023).

11. Adeyeye, S., Ashaolu, T., Bolaji, O., Abegunde, T. & Omoyajowo, A. Africa and the nexus of poverty, malnutrition, and diseases. *Critical Rev. Food Sci. Nutr.* **63**, 641–656, DOI: 10.1080/10408398.2021.1952160 (2021).

12. Ramsay, A. & Kamau, E. Neglected tropical diseases remain a considerable public health challenge in west africa. *Trop. Medicine Infect. Dis.* **10**, 77, DOI: 10.3390/tropicalmed10030077 (2025).

13. George, N., David, S., Nabiryo, M. *et al.* Addressing neglected tropical diseases in africa: a health equity perspective. *Glob. Heal. Res. Policy* **8**, 30, DOI: 10.1186/s41256-023-00314-1 (2023).

14. Chotun, N., Eaton, J., Anagbogu, I. *et al.* Sustaining success through strategies for post-elimination management of neglected tropical diseases in african union member states. *Front. Trop. Dis.* **5**, 1421522, DOI: 10.3389/fitd.2024.1421522 (2024).

15. Muñoz-Zanzi, C., Dreyfus, A., Limothai, U. *et al.* Leptospirosis—improving healthcare outcomes for a neglected tropical disease. *Open Forum Infect. Dis.* **12**, DOI: 10.1093/ofid/ofaf035 (2025).

16. Brattig, N., Bergquist, R., Vienneau, D. *et al.* Geography and health: role of human translocation and access to care. *Infect. Dis. Poverty* **13**, 37, DOI: 10.1186/s40249-024-01205-4 (2024).

17. Adams, A., Vuckovic, M., Birch, E. *et al.* Eliminating neglected tropical diseases in urban areas: A review of challenges, strategies and research directions for successful mass drug administration. *Trop. Medicine Infect. Dis.* **3**, 122, DOI: 10.3390/tropicalmed3040122 (2018).

18. Brindley, P. J., Hotez, P. J. & Kamhawi, S. Revisiting what constitutes a neglected tropical disease? *PLOS Neglected Trop. Dis.* **19**, e0012794 (2025).

19. Marcondes, M. *et al.* Knowledge and practices on lymphatic filariasis in patients with filarial morbidity in an endemic area in northeastern brazil. *Front. Trop. Dis.* **5**, 1368589 (2024).

20. Abdul Rahman, N., Rajaratnam, V., Burchell, G. L., Peters, R. M. & Zweekhorst, M. B. Experiences of living with leprosy: A systematic review and qualitative evidence synthesis. *PLoS neglected tropical diseases* **16**, e0010761 (2022).

21. Frallonardo, L. *et al.* Onchocerciasis: Current knowledge and future goals. *Front. Trop. Dis.* **3**, 986884 (2022).

22. Yotsu, R. R. *et al.* Buruli ulcer: a review of the current knowledge. *Curr. tropical medicine reports* **5**, 247–256 (2018).

23. Alderton, D. L., Ackley, C. & Trueba, M. L. The psychosocial impacts of skin-neglected tropical diseases (sntds) as perceived by the affected persons: A systematic review. *PLOS Neglected Trop. Dis.* **18**, e0012391 (2024).

24. Jayashankar, C. *et al.* Neglected tropical diseases: A comprehensive review. *Cureus* **16** (2024).

25. Houweling, T. A. *et al.* Socioeconomic inequalities in neglected tropical diseases: a systematic review. *PLoS neglected tropical diseases* **10**, e0004546 (2016).

26. Hofstraat, K. & van Brakel, W. H. Social stigma towards neglected tropical diseases: a systematic review. *Int. health* **8**, i53–i70 (2016).

27. Mahapatra, B. *et al.* Economic evaluations of neglected tropical disease interventions in low-and middle-income countries: a systematic review protocol. *JBI Evid. Synth.* **22**, 1582–1593 (2024).

28. World Health Organization. Global report on neglected tropical diseases 2024: Executive summary. Tech. Rep., World Health Organization, Geneva (2024). Accessed September 2025.

29. Al-Tameemi, K. & Kabakli, R. Lymphatic filariasis: an overview. *Asian J. Pharm. Clin. Res.* **12**, 1–5, DOI: 10.22159/ajpcr.2019.v12i12.35646 (2019).

30. Al-Tameemi, K. & Kabakli, R. Lymphatic filariasis: an overview. *Asian J Pharm Clin Res* **12**, 1–5 (2019).

31. Chami, G., Fenwick, A., Bulte, E. *et al.* Influence of schistosoma mansoni and hookworm infection intensities on anaemia in ugandan villages. *PLoS Neglected Trop. Dis.* **9**, e0004193, DOI: 10.1371/journal.pntd.0004193 (2015).

32. Chami, G. F. *et al.* Influence of schistosoma mansoni and hookworm infection intensities on anaemia in ugandan villages. *PLoS neglected tropical diseases* **9**, e0004193 (2015).

33. Al-Haidari, S., Mahdy, M., Al-Mekhlafi, A. *et al.* Intestinal schistosomiasis among schoolchildren in sana'a governorate, yemen: Prevalence, associated factors and its effect on nutritional status and anemia. *PLoS Neglected Trop. Dis.* **15**, e0009757, DOI: 10.1371/journal.pntd.0009757 (2021).

34. Al-Haidari, S. A. *et al.* Intestinal schistosomiasis among schoolchildren in sana'a governorate, yemen: Prevalence, associated factors and its effect on nutritional status and anemia. *PLoS Neglected Trop. Dis.* **15**, e0009757 (2021).

35. Otabil, K., Basáñez, M., Ankrah, B. *et al.* Persistence of onchocerciasis and associated dermatologic and ophthalmic pathologies after 27 years of ivermectin mass drug administration in the middle belt of ghana. *Trop. Medicine & Int. Heal.* **28**, 844–854, DOI: 10.1111/tmi.13937 (2023).

36. Otabil, K. B. *et al.* Persistence of onchocerciasis and associated dermatologic and ophthalmic pathologies after 27 years of ivermectin mass drug administration in the middle belt of ghana. *Trop. Medicine & Int. Heal.* **28**, 844–854 (2023).

37. Silva, M., Andrade, M., de Oliveira, Y. *et al.* Trachoma as cause of blindness: literature review. *Int. Arch. Medicine* **10**, DOI: 10.3823/2326 (2017).

38. Silva, M. B. A. *et al.* Trachoma as cause of blindness: literature review. *Int. Arch. Medicine* **10** (2017).

39. Molyneux, D. Neglected tropical diseases. *Community eye health* **26**, 21 (2013).

40. Quansah, E., Sarpong, E. & Karikari, T. K. Disregard of neurological impairments associated with neglected tropical diseases in africa. *Eneurologicalsci* **3**, 11–14 (2016).

41. Kuper, H. Neglected tropical diseases and disability—what is the link? *Transactions The Royal Soc. Trop. Medicine Hyg.* **113**, 839–844 (2019).

42. Nour, N. M. Schistosomiasis: health effects on women. *Rev. Obstet. Gynecol.* **3**, 28 (2010).

43. Wanyonyi, W., Mulambalah, C., Mulama, D. *et al.* Malaria and geohelminthiasis coinfections in expectant women: effect on maternal health and birth outcomes in a malaria endemic region in kenya. *J. Parasitol. Res.* **2018**, 2613484, DOI: 10.1155/2018/2613484 (2018).

44. Wanyonyi, W. A., Mulambalah, C. S., Mulama, D. H., Omukunda, E. & Siteti, D. I. Malaria and geohelminthiasis coinfections in expectant women: effect on maternal health and birth outcomes in a malaria endemic region in kenya. *J. Parasitol. Res.* **2018**, 2613484 (2018).

45. Watts, C. Neglected tropical diseases: a dfid perspective. *PLoS neglected tropical diseases* **11**, e0005492 (2017).

46. Ochola, E., Karanja, D. & Elliott, S. The impact of neglected tropical diseases (ntds) on health and wellbeing in sub-saharan africa (ssa): A case study of kenya. *PLoS Neglected Trop. Dis.* **15**, e0009131, DOI: 10.1371/journal.pntd.0009131 (2021).

47. Norris, J., Adelman, C., Spantchak, Y. & Marano, K. Social and economic impact review on neglected tropical diseases. *Econ. Policy/Briefing Pap. Wash. DC: Hudson Inst.* (2012).

48. Immurana, M. *et al.* The effects of selected neglected tropical diseases on economic performance at the macrolevel in africa. *BMC Infect. Dis.* **24**, 462 (2024).

49. Ochola, E. A., Karanja, D. M. & Elliott, S. J. The impact of neglected tropical diseases (ntds) on health and wellbeing in sub-saharan africa (ssa): A case study of kenya. *PLOS Neglected tropical diseases* **15**, e0009131 (2021).

50. Koschorke, M., Al-Haboubi, Y. H., Tseng, P.-C., Semrau, M. & Eaton, J. Mental health, stigma, and neglected tropical diseases: A review and systematic mapping of the evidence. *Front. Trop. Dis.* **3**, 808955 (2022).

51. Hamill, L. C. *et al.* People are neglected, not diseases: the relationship between disability and neglected tropical diseases. *Transactions The Royal Soc. Trop. Medicine Hyg.* **113**, 829–834 (2019).

52. Molyneux, D., Dean, L., Adekeye, O. *et al.* The changing global landscape of health and disease: Addressing challenges and opportunities for sustaining progress towards control and elimination of neglected tropical diseases (ntds). *Parasitology* **145**, 1647–1654, DOI: 10.1017/S0031182018000069 (2018).

53. Forbes, K., Basáñez, M., Hollingsworth, T. & Anderson, R. Introduction to the special issue: Challenges and opportunities in the fight against neglected tropical diseases: A decade from the london declaration on ntds. *Philos. Transactions Royal Soc. B: Biol. Sci.* **378**, DOI: 10.1098/rstb.2022.0272 (2023).

54. Bharadwaj, M., Bengtson, M., Golverdingen, M., Waling, L. & Dekker, C. Diagnosing point-of-care diagnostics for neglected tropical diseases. *Adv. Med. Imaging, Detect. Diagn.* 907–934 (2023).

55. Aborode, A. T. *et al.* Bringing lab to the field: Exploring innovations in point-of-care diagnostics for the rapid detection and management of tropical diseases in resource-limited settings. *Adv. Biomark. Sci. Technol.* (2025).

56. Weigl, B. H., Neogi, T. & McGuire, H. Point-of-care diagnostics in low-resource settings and their impact on care in the age of the noncommunicable and chronic disease epidemic. *J. laboratory automation* **19**, 248–257 (2014).

57. Deroco, P. B., Junior, D. W. & Kubota, L. T. Recent advances in point-of-care biosensors for the diagnosis of neglected tropical diseases. *Sensors actuators B: chemical* **349**, 130821 (2021).

58. Kabbas-Piñango, E. *et al.* Reproducibility matters: intra-and inter-sample variation of the point-of-care circulating cathodic antigen test in two schistosoma mansoni endemic areas in uganda. *Philos. Transactions Royal Soc. B* **378**, 20220275 (2023).

59. Rotejanaprasert, C. *et al.* Evaluation of kato-katz and multiplex quantitative polymerase chain reaction performance for clinical helminth infections in thailand using a latent class analysis. *Philos. Transactions Royal Soc. B* **378**, 20220281 (2023).

60. Makau-Barasa, L. K. *et al.* Systems thinking in the prevention, control, and elimination of neglected tropical diseases (ntds). *Int. J. Infect. Dis.* **152**, 107810 (2025).

61. Hietanen, H., Pfavayi, L. T. & Mutapi, F. Unlocking the blueprint to eliminating neglected tropical diseases: A review of efforts in 50 countries that have eliminated at least 1 ntd. *PLOS Neglected Trop. Dis.* **19**, e0013424 (2025).

62. WHO. Reported number of people requiring interventions against neglected tropical diseases (ntds) (2023).

63. Anderson, R., Turner, H., Farrell, S. & Truscott, J. Studies of the transmission dynamics, mathematical model development and the control of schistosome parasites by mass drug administration in human communities. *Adv. parasitology* **94**, 199–246 (2016).

64. Modi, A., Vaishnav, K., Kothiya, K. & Alexander, N. Lymphatic filariasis elimination endgame in an urban indian setting: The roles of surveillance and residual microfilaremia after mass drug administration. *Infect. Dis. Poverty* **10**, 73, DOI: 10.1186/s40249-021-00856-x (2021).

65. Rotejanaprasert, C., Chuaicharoen, P., Prada, J. *et al.* Evaluation of kato-katz and multiplex quantitative polymerase chain reaction performance for clinical helminth infections in thailand using a latent class analysis. *Philos. Transactions Royal Soc. B: Biol. Sci.* **378**, DOI: 10.1098/rstb.2022.0281 (2023).

66. Branda, F. *et al.* Assessing the burden of neglected tropical diseases in low-income communities: challenges and solutions. *Viruses* **17**, 29 (2024).

67. Melo, G. B. T. *et al.* Evolution of research funding for neglected tropical diseases in brazil, 2004–2020. *PLoS Neglected Trop. Dis.* **17**, e0011134 (2023).

68. Fall, L. *et al.* Factors associated with the knowledge, attitudes, and practices of primary healthcare workers regarding neglected tropical diseases with skin manifestations in the dakar region, senegal, 2022. *Trop. Medicine Infect. Dis.* **9**, DOI: 10.3390/tropicalmed9110267 (2024).

69. Sangare, M. *et al.* Understanding the barriers and facilitators related to never treatment during mass drug administration among mobile and migrant populations in mali: a qualitative exploratory study. *BMJ Glob. Heal.* **9**, e015671 (2024).

70. Durojaye, A. B. *et al.* Enhancing communication strategies in controlling neglected tropical diseases in nigeria. *Public Heal. Challenges* **2**, e100 (2023).

71. Assefa, A., Erko, B., Gundersen, S. G., Medhin, G. & Berhe, N. Low awareness and common misconceptions about schistosomiasis in endemic lowland areas in western ethiopia: a mixed-methods study. *BMC Public Heal.* **21**, 1064 (2021).

72. Klepac, P. *et al.* Climate change, malaria and neglected tropical diseases: a scoping review. *Transactions The Royal Soc. Trop. Medicine Hyg.* **118**, 561–579 (2024).

73. Kelly-Hope, L. A., Harding-Esch, E. M., Willems, J., Ahmed, F. & Sanders, A. M. Conflict-climate-displacement: a cross-sectional ecological study determining the burden, risk and need for strategies for neglected tropical disease programmes in africa. *BMJ open* **13**, e071557 (2023).

74. Anderson, R., Turner, H., Farrell, S. & Truscott, J. Studies of the transmission dynamics, mathematical model development and the control of schistosome parasites by mass drug administration in human communities. In Basáñez, M. G. & Anderson, R. M. (eds.) *Mathematical Models for Neglected Tropical Diseases*, vol. 94 of *Advances in Parasitology*, 199–246, DOI: 10.1016/bs.apar.2016.06.003 (Academic Press, 2016).

75. Page, M. J. *et al.* The prisma 2020 statement: an updated guideline for reporting systematic reviews. *bmj* **372** (2021).

76. Qaiser, A. *et al.* Support vector machine outperforms other machine learning models in early diagnosis of dengue using routine clinical data. *Adv. virology* **2024**, 5588127 (2024).

77. Hu, T.-Y., Chow, J. C., Chien, T.-W. & Chou, W. Detecting dengue fever in children using online rasch analysis to develop algorithms for parents: An app development and usability study. *Medicine* **102**, e33296 (2023).

78. Shafqat, S. *et al.* Leveraging deep learning for designing healthcare analytics heuristic for diagnostics. *Neural processing letters* **55**, 53–79 (2023).

79. Mayrose, H. *et al.* An explainable artificial intelligence integrated system for automatic detection of dengue from images of blood smears using transfer learning. *IEEE Access* **12**, 41750–41762 (2024).

80. Dsilva, L. R. *et al.* Wavelet scattering-and object detection-based computer vision for identifying dengue from peripheral blood microscopy. *Int. J. Imaging Syst. Technol.* **34**, e23020 (2024).

81. Mayrose, H., Bairy, G. M., Sampathila, N., Belurkar, S. & Saravu, K. Machine learning-based detection of dengue from blood smear images utilizing platelet and lymphocyte characteristics. *Diagnostics* **13**, 220 (2023).

82. Rada, L., Kumar, P., Martin-Gonzalez, A. & Brito-Loeza, C. Chagas parasite classification in blood sample images using different machine learning architectures. *Med. & Biol. Eng. & Comput.* **62**, 195–206 (2024).

83. Hevia-Montiel, N., Haro, P., Guillermo-Cordero, L. & Perez-Gonzalez, J. Deep learning–based segmentation of trypanosoma cruzi nests in histopathological images. *Electronics* **12**, 4144 (2023).

84. de Miranda, V. L. *et al.* Cellphone picture-based, genus-level automated identification of chagas disease vectors: Effects of picture orientation on the performance of five machine-learning algorithms. *Ecol. Informatics* **79**, 102430 (2024).

85. Tekle, E. *et al.* Deepleish: a deep learning based support system for the detection of leishmaniasis parasite from giemsa-stained microscope images. *BMC Med. Imaging* **24**, 152 (2024).

86. Sadeghi, A. *et al.* A deep learning-based model for detecting leishmania amastigotes in microscopic slides: a new approach to telemedicine. *BMC Infect. Dis.* **24**, 551 (2024).

87. Borges, A. L. *et al.* Visceral leishmaniasis detection using deep learning techniques and multiple color space bands. In *International Conference on Intelligent Systems Design and Applications*, 492–502 (Springer, 2022).

88. Meulah, B. *et al.* Validation of artificial intelligence-based digital microscopy for automated detection of schistosoma haematobium eggs in urine in gabon. *PLoS neglected tropical diseases* **18**, e0011967 (2024).

89. Oliveira, B. A. S. *et al.* Automated diagnosis of schistosomiasis by using faster r-cnn for egg detection in microscopy images prepared by the kato–katz technique. *Neural Comput. Appl.* **34**, 9025–9042 (2022).

90. Guo, Z. *et al.* Feasibility of ultrasound radiomics based models for classification of liver fibrosis due to schistosoma japonicum infection. *PLOS Neglected Trop. Dis.* **18**, e0012235 (2024).

91. Saha, A. *et al.* Deep humoral profiling coupled to interpretable machine learning unveils diagnostic markers and pathophysiology of schistosomiasis. *Sci. translational medicine* **16**, eadk7832 (2024).

92. Ebersbach, J. C. *et al.* Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for differential identification of adult schistosoma worms. *Parasites & Vectors* **16**, 20 (2023).

93. Acula, D. D., Buico, C., Cruzada, G., Encelan, M. & Rivera, C. Soil transmitted helminth egg detection and classification in fecal smear images using faster region-based convolutional neural network with residual network-50. In *International Conference on Green Energy, Computing and Intelligent Technology (GEn-CITy 2023)*, vol. 2023, 284–291 (IET, 2023).

94. Milad, D., Antaki, F., Robert, M.-C. & Duval, R. Development and deployment of a smartphone application for diagnosing trachoma: leveraging code-free deep learning and edge artificial intelligence. *Saudi J. Ophthalmol.* **37**, 200–206 (2023).

95. Zewudie, M. S., Xiong, S., Yu, X., Wu, X. & Mehamed, M. A. Adaptive feature selection for active trachoma image classification. *Knowledge-Based Syst.* **294**, 111764 (2024).

96. De Souza, M. L. M., Lopes, G. A., Branco, A. C., Fairley, J. K. & Fraga, L. A. D. O. Leprosy screening based on artificial intelligence: development of a cross-platform app. *JMIR mHealth uHealth* **9**, e23718 (2021).

97. Barbieri, R. R. *et al.* Reimagining leprosy elimination with ai analysis of a combination of skin lesion images with demographic and clinical data. *The Lancet Reg. Heal.* **9** (2022).

98. Keshavamurthy, R. *et al.* Machine learning to improve the understanding of rabies epidemiology in low surveillance settings. *Sci. Reports* **14**, 25851 (2024).

99. Punyapornwithaya, V. *et al.* Time series analysis and forecasting of the number of canine rabies confirmed cases in thailand based on national-level surveillance data. *Front. Vet. Sci.* **10**, 1294049 (2023).

100. Zhang, X. *et al.* Precision mapping of snail habitat in lake and marshland areas: Integrating environmental and textural indicators using random forest modeling. *Heliyon* **10** (2024).

101. Xu, N. *et al.* Prediction of oncomelania hupensis distribution in association with climate change using machine learning models. *Parasites & Vectors* **16**, 377 (2023).

102. Yang, J., Mosabbir, A. A., Raheem, E., Hu, W. & Hossain, M. S. Demographic characteristics, clinical symptoms, biochemical markers and probability of occurrence of severe dengue: A multicenter hospital-based study in bangladesh. *PLoS Neglected Trop. Dis.* **17**, e0011161 (2023).

103. Nandwani, S., Bhakhri, B. K., Singh, N., Rai, R. & Singh, D. K. Early hematological parameters as predictors for outcomes in children with dengue in northern india: A retrospective analysis. *Revista da Sociedade Brasileira de Medicina Trop.* **54**, e05192020 (2021).

104. Li, G. *et al.* Comparison of three data mining models for prediction of advanced schistosomiasis prognosis in the hubei province. *PLoS neglected tropical diseases* **12**, e0006262 (2018).

105. Bamorovat, M. *et al.* A novel diagnostic and prognostic approach for unresponsive patients with anthroponotic cutaneous leishmaniasis using artificial neural networks. *PLoS One* **16**, e0250904 (2021).

106. Gerardo, C. J. *et al.* Association of a network of immunologic response and clinical features with the functional recovery from crotalinae snakebite envenoming. *Front. immunology* **12**, 628113 (2021).

107. Gaye, P. M. *et al.* Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry traces the geographical source of biomphalaria pfeifferi and bulinus forskalii, involved in schistosomiasis transmission. *Infect. Dis. Poverty* **13**, 11 (2024).

108. Xu, J. *et al.* Forecast of dengue cases in 20 chinese cities based on the deep learning method. *Int. journal environmental research public health* **17**, 453 (2020).

109. Manocha, A., Bhatia, M. & Kumar, G. Smart monitoring solution for dengue infection control: A digital twin-inspired approach. *Comput. Methods Programs Biomed.* **257**, 108459 (2024).

110. Nanda, A. K. *et al.* Forecasting deep learning-based risk assessment of vector-borne diseases using hybrid methodology. *Technol. Heal. Care* **32**, 3341–3361 (2024).

111. de Andrade Rodrigues, R. S. *et al.* Prediction of the occurrence of leprosy reactions based on bayesian networks. *Front. Medicine* **10**, 1233220 (2023).

112. Lagatie, O. *et al.* Multimodal biomarker discovery for active onchocerca volvulus infection. *PLoS Neglected Trop. Dis.* **15**, e0009999 (2021).

113. Villamizar-Monsalve, M. A. *et al.* Development and application of an in vitro drug screening assay for schistosoma mansoni schistosomula using yolov5. *Biomedicines* **12**, 2894 (2024).

114. Zeng, X. *et al.* Gnn-ddas: Drug discovery for identifying anti-schistosome small molecules based on graph neural network. *J. Comput. Chem.* **45**, 2825–2834 (2024).

115. Santos, E. S. d. A. *et al.* Deep multitask learning-driven discovery of new compounds targeting leishmania infantum. *ACS omega* **9**, 51271–51284 (2024).

116. Melo-Filho, C. C. *et al.* Qsar-driven discovery of novel chemical scaffolds active against schistosoma mansoni. *J. chemical information modeling* **56**, 1357–1372 (2016).

117. Francisco, M. E., Carvajal, T. M. & Watanabe, K. Hybrid machine learning approach to zero-inflated data improves accuracy of dengue prediction. *PLOS Neglected Trop. Dis.* **18**, e0012599 (2024).

118. Patil, S. & Pandya, S. Forecasting dengue hotspots associated with variation in meteorological parameters using regression and time series models. *Front. public health* **9**, 798034 (2021).

119. Akter, T., Islam, M. T., Hossain, M. F. & Ullah, M. S. A comparative study between time series and machine learning technique to predict dengue fever in dhaka city. *Discret. Dyn. Nat. Soc.* **2024**, 2757381 (2024).

120. Sebastianelli, A. *et al.* A reproducible ensemble machine learning approach to forecast dengue outbreaks. *Sci. reports* **14**, 3807 (2024).

121. Nguyen, V.-H. *et al.* Deep learning models for forecasting dengue fever based on climate data in vietnam. *PLoS Neglected Trop. Dis.* **16**, e0010509 (2022).

122. Panja, M. *et al.* An ensemble neural network approach to forecast dengue outbreak based on climatic condition. *Chaos, Solitons & Fractals* **167**, 113124 (2023).

123. Tuan, D. A. Leveraging climate data for dengue forecasting in ba ria vung tau province, vietnam: An advanced machine learning approach. *Trop. Medicine Infect. Dis.* **9**, 250 (2024).

124. Barboza, L. A. *et al.* Assessing dengue fever risk in costa rica by using climate variables and machine learning techniques. *PLOS Neglected Trop. Dis.* **17**, e0011047 (2023).

125. Gupta, G. *et al.* Ddpm: A dengue disease prediction and diagnosis model using sentiment analysis and machine learning algorithms. *Diagnostics* **13**, 1093 (2023).

126. Husin, N. A. *et al.* Performance of hybrid gann in comparison with other standalone models on dengue outbreak prediction. *J. Comput. Sci.* **12**, 300–306 (2016).

127. Suggala, R. K., Krishna, M. V. & Swain, S. K. Discover the new factor for dengue fever outbreaks and predicted using bayes network-pso (bn-pso). *J. Inf. Sci. & Eng.* **39** (2023).

128. Singer, B. J. *et al.* Development of prediction models to identify hotspots of schistosomiasis in endemic regions to guide mass drug administration. *Proc. Natl. Acad. Sci.* **121**, e2315463120 (2024).

129. Su, Q. *et al.* Unraveling trends in schistosomiasis: deep learning insights into national control programs in china. *Epidemiol. Heal.* **46**, e2024039 (2024).

130. Ahamjik, I. *et al.* Contribution of artificial intelligence for understanding animal rabies epidemiology in morocco: What are the perspectives of an innovative and predictive approaches? *One Heal.* **19**, 100874 (2024).

131. Shah, J., Khan, H., Ismail, E. A., Awaad, F. A. & Kumar, A. Modeling scabies transmission dynamics: a stochastic approach with spectral collocation and neural network insights. *The Eur. Phys. J. Plus* **140**, 66 (2025).

132. Majeed, M. A., Shafri, H. Z. M., Zulkafli, Z. & Wayayok, A. A deep learning approach for dengue fever prediction in malaysia using lstm with spatial attention. *Int. journal environmental research public health* **20**, 4130 (2023).

133. DN, A. N., Li, Y., Hsu, C.-Y. *et al.* Artificial intelligence approach for severe dengue early warning system. *Stud. health technology informatics* **310**, 881–885 (2024).

134. Machuca, J. V. B., Díaz, D. H. S. & Serra, C. E. S. Cluster-based lstm models to improve dengue cases forecast. *CLEI Electron. J.* **26**, 4–1 (2023).

135. Ullah, M. A., Mim, A. S., Hasan, M. N. & Sadik, M. R. Deep learning based forecasting models of dengue outbreak in bangladesh: Comparative analysis of lstm, rnn, and gru models using multivariate variables with a two-decade dataset. In *2024 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)*, 1–6 (IEEE, 2024).

136. Doni, A. R. & Sasipraba, T. Lstm-rnn based approach for prediction of dengue cases in india. *Ingénierie des Systèmes d'Information* **25** (2020).

137. Islam, M. S., Shahrear, P., Saha, G., Ataullha, M. & Rahman, M. S. Mathematical analysis and prediction of future outbreak of dengue on time-varying contact rate using machine learning approach. *Comput. biology medicine* **178**, 108707 (2024).

138. Laureano-Rosario, A. E. *et al.* Application of artificial neural networks for dengue fever outbreak predictions in the northwest coast of yucatan, mexico and san juan, puerto rico. *Trop. Medicine Infect. Dis.* **3**, 5 (2018).

139. Khan, M. A. R., Akter, J., Ahammad, I., Ejaz, S. & Jaman Khan, T. Dengue outbreaks prediction in bangladesh perspective using distinct multilayer perceptron nn and decision tree. *Heal. Inf. Sci. Syst.* **10**, 32 (2022).

140. Glidden, C. K. *et al.* Climate and urbanization drive changes in the habitat suitability of schistosoma mansoni competent snails in brazil. *Nat. Commun.* **15**, 4838 (2024).

141. Sarma, D. *et al.* Dengue prediction using machine learning algorithms. In *2020 IEEE 8th R10 humanitarian technology conference (R10-HTC)*, 1–6 (IEEE, 2020).

142. Iqbal, N. & Islam, M. Machine learning for dengue outbreak prediction: A performance evaluation of different prominent classifiers. *Informatica* **43** (2019).

143. Agarwala, S. *et al.* Precision death forecasting of dengue outbreaks in bangladesh by blending statistical and machine learning models. In *2024 IEEE International Conference on Power, Electrical, Electronics and Industrial Applications (PEEIACON)*, 981–986 (IEEE, 2024).

144. Aleixo, R., Kon, F., Rocha, R., Camargo, M. S. & De Camargo, R. Y. Predicting dengue outbreaks with explainable machine learning. In *2022 22nd IEEE International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, 940–947 (IEEE, 2022).

145. Sarwar, M. T. & Al Mamun, M. Prediction of dengue using machine learning algorithms: Case study dhaka. In *2022 4th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)*, 1–6 (IEEE, 2022).

146. Mai, S. T. *et al.* Dengue fever: From extreme climates to outbreak prediction. In *2022 IEEE International Conference on Data Mining (ICDM)*, 1083–1088 (IEEE, 2022).

147. Halim, S., Octavia, T., Handojo, A. *et al.* Dengue fever outbreak prediction in surabaya using a geographically weighted regression. In *2019 4th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON)*, 1–5 (IEEE, 2019).

148. Anggraeni, W. *et al.* Prediction of dengue fever outbreak based on climate factors using fuzzy-logistic regression. In *2020 international seminar on intelligent technology and its applications (ISITIA)*, 199–204 (IEEE, 2020).

149. Zheng, J.-X. *et al.* Infestation risk of the intermediate snail host of schistosoma japonicum in the yangtze river basin: improved results by spatial reassessment and a random forest approach. *Infect. diseases poverty* **10**, 34–46 (2021).

150. Somboonsak, P. Forecasting dengue fever epidemics using arima model. In *Proceedings of the 2019 2nd Artificial Intelligence and Cloud Computing Conference*, 144–150 (2019).

151. Rajan, S. D. Integrated machine learning paradigms with simulation based mathematical modeling for dengue fever. In *2024 International Conference on Social and Sustainable Innovations in Technology and Engineering (SASI-ITE)*, 131–136 (IEEE, 2024).

152. Gallini, N., Hancock, R. E., Lester, R., Sharp, B. & Wasan, K. M. Implementation and impact of the global access principles at the university of british columbia: current successes and future challenges. *Front. Pharmacol.* **14**, 1271618 (2023).

153. Li, Q. *et al.* Optimized strategy for schistosomiasis elimination: results from marginal benefit modeling. *Parasites & Vectors* **16**, 419 (2023).

154. Whitson, C. C. *et al.* Photographic grading for trachoma diagnosis within trachoma impact surveys in amhara region, ethiopia. *Transactions The Royal Soc. Trop. Medicine Hyg.* **117**, 111–117 (2023).

155. Aribodor, O. B. *et al.* Profiling the knowledge of female medical/para-medical students, and expertise of health care professionals on female genital schistosomiasis in anambra, south eastern nigeria. *PLOS Neglected Trop. Dis.* **17**, e0011132 (2023).

156. Chang, Y.-C., Chiu, Y.-W. & Chuang, T.-W. Linguistic pattern–infused dual-channel bidirectional long short-term memory with attention for dengue case summary generation from the program for monitoring emerging diseases–mail database: algorithm development study. *JMIR public health surveillance* **8**, e34583 (2022).

157. Kamarudin, A. N. A., Zainol, Z. & Kassim, N. F. A. Predicting the misconception of dengue disease based on the awareness survey. *Bull. Electr. Eng. Informatics* **9**, 1631–1637 (2020).

158. Purnomo, M. H., Fahruzi, I. & Anggraeni, W. Enhance computing performance through deep learning and blockchain integration for public health problems. In *2020 3rd International Conference on Biomedical Engineering (IBIOMED)*, 65–70 (IEEE, 2020).

159. Eneanya, O. A. *et al.* Mapping the baseline prevalence of lymphatic filariasis across nigeria. *Parasites & vectors* **12**, 440 (2019).

160. Maia-Elkhoury, A. N., OB Valadas, S. Y., Puppim-Buzanovsky, L., Rocha, F. & Sanchez-Vazquez, M. J. Sisleish: A multi-country standardized information system to monitor the status of leishmaniasis in the americas. *PLoS Neglected Trop. Dis.* **11**, e0005868 (2017).

161. da Silva Neto, S. R. *et al.* Machine learning and deep learning techniques to support clinical diagnosis of arboviral diseases: A systematic review. *PLoS neglected tropical diseases* **16**, e0010061 (2022).

162. Attai, K., Amannejad, Y., Vahdat Pour, M., Obot, O. & Uzoka, F.-M. A systematic review of applications of machine learning and other soft computing techniques for the diagnosis of tropical diseases. *Trop. Medicine Infect. Dis.* **7**, 398 (2022).

163. Kumar, S., Arif, T., Alotaibi, A. S., Malik, M. B. & Manhas, J. Advances towards automatic detection and classification of parasites microscopic images using deep convolutional neural network: methods, models and research directions. *Arch. Comput. Methods Eng.* **30**, 2013–2039 (2023).

164. Siriyasatien, P., Chadsuthi, S., Jampachaisri, K. & Kesorn, K. Dengue epidemics prediction: A survey of the state-of-the-art based on data science processes. *IEEE Access* **6**, 53757–53795 (2018).

165. Leung, X. Y. *et al.* A systematic review of dengue outbreak prediction models: Current scenario and future directions. *PLoS neglected tropical diseases* **17**, e0010631 (2023).

166. Khew, C., Akbar, R. & Mohd-Assaad, N. Progress and challenges for the application of machine learning for neglected tropical diseases. *F1000Research* **12**, 287 (2025).

167. Messina, J. P. *et al.* A global compendium of human dengue virus occurrence. *Sci. Data* **1**, 1–6 (2014).

168. Chaudhury, S. *et al.* Dengue virus antibody database: Systematically linking serotype-specificity with epitope mapping in dengue virus. *PLoS neglected tropical diseases* **11**, e0005395 (2017).

169. Hatcher, E. L. *et al.* Virus variation resource–improved response to emergent viral outbreaks. *Nucleic acids research* **45**, D482–D490 (2017).

170. Kaur, I., Sandhu, A. K. & Kumar, Y. Artificial intelligence techniques for predictive modeling of vector-borne diseases and its pathogens: a systematic review. *Arch. Comput. Methods Eng.* **29**, 3741–3771 (2022).

171. Cecchi, G. *et al.* Disease intelligence and modelling for progressive control of animal trypanosomosis in africa. *Trends Parasitol.* (2025).

172. Abadi, M. *et al.* {TensorFlow}: a system for {Large-Scale} machine learning. In *12th USENIX symposium on operating systems design and implementation (OSDI 16)*, 265–283 (2016).

173. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. *Adv. neural information processing systems* **32** (2019).

174. Pedregosa, F. *et al.* Scikit-learn: Machine learning in python. *J. machine Learn. research* **12**, 2825–2830 (2011).

175. Dehnavieh, R., Haghdoost, A., Khosravi, A. *et al.* The district health information system (dhis2): A literature review and meta-synthesis of its strengths and operational challenges. *Heal. Inf. Manag. J.* **48**, 62–75 (2019).

176. Mohammed-Rajput, N. A., Smith, D. C., Mamlin, B. W. & Biondich, P. Openmrs, a global medical records system collaborative: Factors influencing successful implementation for the open mrs collaborative investigators. In *AMIA Annual Symposium Proceedings*, 960–968 (2011).

177. Hartung, C. *et al.* Open data kit: tools to build information services for developing regions. In *Proceedings of the 4th ACM/IEEE international conference on information and communication technologies and development*, 1–12 (2010).

178. Gorelick, N., Hancher, M., Dixon, M. *et al.* Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote. Sens. Environ.* **202**, 18–27 (2017).

179. Nagappan, E. & Egambaram, S. Quantum geographic information system (qgis) for mapping of study areas. *Biotica Res. Today* **5**, 70–72 (2023).

180. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Adv. Neural Inf. Process. Syst.* **30**, 4765–4774 (2017).

181. Ribeiro, M. T., Singh, S. & Guestrin, C. "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144 (2016).

182. Kokhlikyan, N. *et al.* Captum: A unified and generic model interpretability library for pytorch. *arXiv preprint arXiv:2009.07896* (2020).

183. Fernandez, F.-G. Torchcam: class activation explorer. https://github.com/frgfm/torch-cam (2020).

184. Alber, M. *et al.* innvestigate neural networks! *J. Mach. Learn. Res.* **20**, 1–8 (2019).

185. Cardoso, M. J. *et al.* Monai: An open-source framework for deep learning in healthcare. *arXiv preprint arXiv:2211.02701* (2022).

186. Bird, S. *et al.* Fairlearn: A toolkit for assessing and improving fairness in ai. *Microsoft Res. Tech. Rep.* (2020).

187. Bellamy, R. K. *et al.* Ai fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM J. Res. Dev.* **63**, 4–1 (2019).

188. Kairouz, P. *et al.* Advances and open problems in federated learning. *Foundations Trends Mach. Learn.* **14**, 1–210 (2021).

189. Zhang, F. *et al.* Recent methodological advances in federated learning for healthcare. *Patterns* **5** (2024).

190. Chaddad, A., Wu, Y. & Desrosiers, C. Federated learning for healthcare applications. *IEEE internet things journal* **11**, 7339–7358 (2023).

191. Yang, Q. *et al.* Federated learning: Opportunities and challenges for health applications. *IEEE Transactions on Big Data* **5**, 20–36 (2019).

192. Jing, L. & Tian, Y. Self-supervised visual feature learning with deep neural networks: A survey. *IEEE transactions on pattern analysis machine intelligence* **43**, 4037–4058 (2020).

193. Chen, T., Kornblith, S., Norouzi, M. & Hinton, G. A simple framework for contrastive learning of visual representations. In *International conference on machine learning*, 1597–1607 (PmLR, 2020).

194. Shurrab, S. & Duwairi, R. Self-supervised learning methods and applications in medical imaging analysis: A survey. *PeerJ Comput. Sci.* **8**, e1045 (2022).

195. Krishnan, R., Rajpurkar, P. & Topol, E. J. Self-supervised learning in medicine and healthcare. *Nat. Biomed. Eng.* **6**, 1346–1352 (2022).

196. Huang, S.-C. *et al.* Self-supervised learning for medical image classification: a systematic review and implementation guidelines. *NPJ Digit. Medicine* **6**, 74 (2023).

197. Bommasani, R. *et al.* On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258* (2021).

198. Gu, Y. *et al.* Domain-specific language model pretraining for biomedical natural language processing. *ACM Transactions on Comput. for Healthc. (HEALTH)* **3**, 1–23 (2021).

199. Achiam, J. *et al.* Gpt-4 technical report. *arXiv preprint arXiv:2303.08774* (2023).

200. Comanici, G. *et al.* Gemini 2.5: Pushing the frontier with advanced reasoning, multimodality, long context, and next generation agentic capabilities. *arXiv preprint arXiv:2507.06261* (2025).

201. Rasmy, L., Xiang, Y., Xie, Z., Tao, C. & Zhi, D. Med-bert: pretrained contextualized embeddings on large-scale structured electronic health records for disease prediction. *NPJ digital medicine* **4**, 86 (2021).

202. Wang, S., Hu, M., Li, Q., Safari, M. & Yang, X. Capabilities of gpt-5 on multimodal medical reasoning. *arXiv preprint arXiv:2508.08224* (2025).

203. Ryu, J. S., Kang, H., Chu, Y. & Yang, S. Vision-language foundation models for medical imaging: a review of current practices and innovations. *Biomed. Eng. Lett.* 1–22 (2025).

204. Wang, Z. *et al.* A foundation model for human-ai collaboration in medical literature mining. *Nat. communications* **16**, 8361 (2025).

205. Chen, Z., Wei, L. & Gao, G. Foundation models for bioinformatics. *Quant. Biol.* **12**, 339–344 (2024).

206. Rillig, M. C., Ågerstrand, M., Bi, M., Gould, K. A. & Sauerland, U. Risks and benefits of large language models for the environment. *Environ. science & technology* **57**, 3464–3466 (2023).

207. Liu, V. & Yin, Y. Green ai: exploring carbon footprints, mitigation strategies, and trade offs in large language model training. *Discov. Artif. Intell.* **4**, 49 (2024).

208. Luccioni, A. S., Viguier, S. & Ligozat, A.-L. Estimating the carbon footprint of bloom, a 176b parameter language model. *J. machine learning research* **24**, 1–15 (2023).

209. Truhn, D., Müller-Franzes, G. & Kather, J. N. The ecological footprint of medical ai. *Eur. Radiol.* **34**, 1176–1178 (2024).

210. Schölkopf, B. *et al.* Toward causal representation learning. *Proc. IEEE* **109**, 612–634 (2021).

211. Pearl, J. *Causality* (Cambridge university press, 2009).

212. Schölkopf, B. *et al.* Causal representation learning. *arXiv preprint arXiv:2106.15661* (2021).

213. Baltrušaitis, T., Ahuja, C. & Morency, L.-P. Multimodal machine learning: A survey and taxonomy. *IEEE transactions on pattern analysis machine intelligence* **41**, 423–443 (2018).

214. Ngiam, J. *et al.* Multimodal deep learning. In *ICML*, vol. 11, 689–696 (2011).

215. Stahlschmidt, S. R., Ulfenborg, B. & Synnergren, J. Multimodal deep learning for biomedical data fusion: a review. *Briefings bioinformatics* **23**, bbab569 (2022).

216. Luo, Y. *et al.* Toward unified ai drug discovery with multimodal knowledge. *Heal. Data Sci.* **4**, 0113 (2024).

217. Barnowska, E. J. *et al.* Diagnosing skin neglected tropical diseases with the aid of digital health tools: A scoping review. *PLOS Digit. Heal.* **3**, e0000629 (2024).

218. van de Sande, W. W. *et al.* The mycetoma knowledge gap: identification of research priorities. *PLoS neglected tropical diseases* **8**, e2667 (2014).

219. Yang, J., Soltan, A. A. & Clifton, D. A. Machine learning generalizability across healthcare settings: insights from multi-site covid-19 screening. *NPJ digital medicine* **5**, 69 (2022).

220. Obermeyer, Z., Powers, B., Vogeli, C. & Mullainathan, S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* **366**, 447–453 (2019).

221. Joseph, J. Algorithmic bias in public health ai: a silent threat to equity in low-resource settings. *Front. Public Heal.* **13**, 1643180 (2025).

222. Schwalbe, N. & Wahl, B. Artificial intelligence and the future of global health. *The Lancet* **395**, 1579–1586 (2020).

223. Hopkins, A. D. Neglected tropical diseases in africa: a new paradigm. *Int. health* **8**, i28–i33 (2016).

224. Hotez, P. J. *Forgotten people, forgotten diseases: the neglected tropical diseases and their impact on global health and development* (John Wiley & Sons, 2021).

225. Mitjà, O. *et al.* Integrated control and management of neglected tropical skin diseases. *PLoS neglected tropical diseases* **11**, e0005136 (2017).

226. Bhatia, M., Chaudhary, J., Jain, A. & Dhingra, A. Current therapeutic strategies for buruli ulcer. In *Emerging Approaches to Tackle Neglected Diseases: From Molecule to End Product*, 166–181 (Bentham Science Publishers, 2024).

227. Rieke, N. *et al.* The future of digital health with federated learning. *NPJ digital medicine* **3**, 119 (2020).

228. Liu, Y. & Zhang, L. Federated learning for healthcare: Opportunities and challenges. *IEEE J. Biomed. Heal. Informatics* **25**, 4153–4160 (2021).

229. Ciecierski-Holmes, T., Singh, R., Axt, M., Brenner, S. & Barteit, S. Artificial intelligence for strengthening healthcare systems in low-and middle-income countries: a systematic scoping review. *NPJ digital medicine* **5**, 162 (2022).

230. Mollura, D. J. *et al.* Artificial intelligence in low-and middle-income countries: innovating global health radiology. *Radiology* **297**, 513–520 (2020).

231. Marey, A. *et al.* A perspective on ai implementation in medical imaging in lmics: challenges, priorities, and strategies. *Eur. Radiol.* 1–12 (2025).

232. El Arab, R. A. & Al Moosa, O. A. Systematic review of cost effectiveness and budget impact of artificial intelligence in healthcare. *npj Digit. Medicine* **8**, 548 (2025).

233. Kastrup, N., Holst-Kristensen, A. W. & Valentin, J. B. Landscape and challenges in economic evaluations of artificial intelligence in healthcare: a systematic review of methodology. *BMC Digit. Heal.* **2**, 1–12 (2024).

234. Yu, L. & Zhai, X. Use of artificial intelligence to address health disparities in low-and middle-income countries: a thematic analysis of ethical issues. *Public Heal.* **234**, 77–83 (2024).

235. Han, H. Challenges of reproducible ai in biomedical data science. *BMC Med. Genomics* **18**, 8 (2025).

236. Petersen, E. *et al.* Responsible and regulatory conform machine learning for medicine: a survey of challenges and solutions. *IEEE access* **10**, 58375–58418 (2022).

237. Markus, A. F., Kors, J. A. & Rijnbeek, P. R. The role of explainability in creating trustworthy artificial intelligence for health care: a comprehensive survey of the terminology, design choices, and evaluation strategies. *J. biomedical informatics* **113**, 103655 (2021).

238. Wilkinson, M. D. *et al.* The fair guiding principles for scientific data management and stewardship. *Sci. data* **3**, 1–9 (2016).

239. Makady, A. *et al.* What is real-world data? a review of definitions based on literature and stakeholder interviews. *Value health* **20**, 858–865 (2017).

240. Groth Jensen, L., Mølgaard Hansen, R., Deleuran Terkildsen, M. & Palmhøj Nielsen, C. Iderha: Unlocking health data for better public health decision-making. *Eur. J. Public Heal.* **35**, ckaf161–968 (2025).