Designing AI Algorithms to Suit Local Context

Yunusa G. Muhammed *[1], Tarisiro Matiza *[2], and Charles B. Delahunt *[3]

[1] Gombe State University, Tudun Wada Gombe Nigeria, yunusa.mohammed@gsu.edu.ng
[2] Market Access Africa, Harare, Zimbabwe, tmatiza@marketaccess.africa
[3] Global Health Labs, Inc, Bellevue, WA, USA, delahunt@uw.edu
* Equal contributions

Abstract. Artificial Intelligence holds tremendous promise to transform healthcare, and if we can effectively convert technical advances in AI into robust deployments, then we can vastly improve quality of life for billions of currently underserved people. But a substantial barrier to this opportunity is a disconnect between AI developers and clinical realities: To deploy successfully, AI development must be shaped at every stage by the specifics of the local deployment context. This integration is complex, it's a necessary condition of success, and it falls to AI teams to carry out. However, because this task sits outside the traditional algorithm-centric focus of AI, it is poorly understood. Therefore this paper describes little-discussed but crucial aspects of an algorithm's path to deployment, each directly relevant to AI researchers and illustrated by concrete examples drawn from experiences in the African healthcare context.
We describe how AI is just one part of a much larger healthcare context, and how a core task of the AI team is to understand and translate this context into AI design choices from the very start of a project. We also include a set of actionable steps accessible to any researcher, which can markedly improve the fitness of an algorithm for future deployment.

Keywords: Healthcare · Domain expertise · Artificial Intelligence · Machine Learning.

1 Introduction

AI has the potential to transform health care, especially for people who currently suffer large gaps in care delivery. However, a serious barrier to this opportunity is a disconnect between AI developers and clinical realities [13]. In this paper, we argue that to deploy successfully, AI development must be shaped, at every stage and in multiple ways, by an understanding of the specifics of the local deployment context. This integration of local context into AI work is complex and difficult, and sits outside the traditional algorithm-centric focus of AI for health care. But it can also only be carried out by AI engineers - for example, while clinicians can specify real-world constraints, only AI engineers can translate these constraints into forces that shape an algorithm. Therefore, this integration of local context is an essential part of the technical workload of the AI researcher, just as important to success as architecture work. This expands the definition of technical AI work.

This topic has been addressed previously, for example by [15], which describes an excellent and comprehensive framework that is however necessarily general. In this paper, we extend the discussion by taking an "in the trenches" viewpoint. We describe some little-discussed but crucial aspects of the path to AI deployment, directly relevant to AI researchers, emphasizing how local context impacts even the earliest stages of algorithm development, with concrete examples drawn from experience in the African healthcare landscape. We hope that this approach highlights a crucial working rule of AI development, that it is just one part of a much larger puzzle, and that non-AI aspects of the local context must be integrated into AI work at every stage. Because not all AI researchers are positioned to fully address the factors we list, we also include a set of actionable steps, accessible to any AI researcher, which can markedly improve the fitness of an algorithm for future deployment.

2 Methods

The "local deployment context" is defined here as the clinical use case, location, and circumstances of the proposed deployment, including the needs of the patient population, infrastructure, medical personnel, and business models. In this section we describe some crucial factors of the local context, with examples to illustrate each. Since they all affect algorithm development, they need to be actively incorporated into AI work *from the start* to enable successful deployment:

1. Domain experts
2. Collaborative planning and goal alignment (includes defining AI goals)
3. Data collection
4. Algorithm design
5. Iterative co-design

2.1 Engage local domain experts

This includes people who understand the specifics of the local context in which the AI algorithm will operate - clinicians, patients, policymakers, and researchers. This is a crucial first step because if the AI design does not incorporate the specific needs and constraints of each stakeholder, deployment will falter at some point due to an unaccounted-for factor. For example:

1. In India, revealing the sex of a fetus in ultrasound (US) is illegal [27] due to risk of female feticide. Thus an US application cannot provide sex-specific information about the fetus to patients, and the US devices themselves are strictly controlled. In Africa, by contrast, US images of the baby bring in the crowds, as parents want a glimpse of the new baby, and thus the images boost support of US in ante-natal care [17]. These contrasting local contexts will shape the basic design of AI applications for obstetric US.

2. Dr. Groesbeck Parham has led development of a successful cervical cancer screening program in Zambia [18]. Though an expert on cervical cancer in the USA, on arrival in Zambia he found a radically different clinical picture, and Zambian gynecologists educated him about the local context (e.g. its deadly combination with HIV infections). After developing a screening method on this basis, he found that conditions in outlying villages were radically different than in Lusaka. Village leaders and village clinic personnel educated him as to what could work. When the system deployed in these village clinics, he found that patient uptake depended on creating more trust, which required teams of local women recruited in each village. Thus, for this highly trained and experienced American gynecologist the journey to a successful deployment depended on aligning his skills with ever more local expertise.

2.2 Collaborative planning and goal alignment

Because local experts have vital knowledge about their situation, without which a clinical problem cannot be even defined, much less solved, AI researchers must solicit collaborations with these experts to define problems and goals:

Jointly define use cases with impact While we as AI researchers have ideas as to what AI can do, the local experts know what is needed locally. In the Zambian example above, a central factor was HIV as an accelerant of cervical cancer progression. Both parties have essential knowledge, but the local clinicians' knowledge has priority: applying AI to a real medical problem is difficult but can yield a useful tool, while an algorithm developed from internal AI perspectives is unlikely to have clinical utility.

Develop shared vocabulary The AI field and the medical field each have a body of definitions and assumptions that can radically differ. Crucially, the fields disagree on what constitutes valid evidence. In medicine, the gold standard is a clinical trial, run by an independent party, which compares a therapeutic to the standard-of-care. In the field of AI, the standard is a comparison to other algorithms on benchmark datasets. By the medical field's standards, an AI-style comparison is inadequate to evaluate a diagnostic for safe deployment. So to deploy our algorithm in a clinic, the evidence must satisfy the medical field's rules, not those of AI. The WHO has a valuable document clarifying this distinction and describing the types of evidence required of AI for healthcare [25].

Understand the local context The AI algorithm is only one piece of a complex puzzle, and the specific local context defines many of the other pieces, with which the AI piece must mesh for successful deployment. These pieces include constraints (and opportunities) of infrastructure, workforce, and logistics.

Price Pricing is a major challenge. It must be sustainable for Ministries of Health (MoHs) or private markets to support an AI tool, and no AI algorithm will see deployment if the costs don't pencil out. African countries have limited budgets, especially with the recent withdrawal of U.S. and other funding. Products procured through non-local funding can be abandoned if the service and maintenance costs for equipment are unsustainable or too high for scale-up.

In Africa, most people seek care in public health facilities, a market with limited resources funded by governments, and the private sector market is very small. This strongly affects pricing and procurement mechanisms, which in turn affect basic design choices for algorithms and hardware. Pricing models vary:

(1) Upfront licensing involves one-time purchase of a perpetual license, often with accompanying hardware. However, the high initial investment often requires buy-in from international funding agencies.

(2) Subscriptions (monthly or annual) are used by many CAD vendors. The lower upfront cost for the AI portion is easier for governments to budget, but reducing high hardware cost impacts AI design choices (e.g. should the AI use dedicated scanning microscopes, or cellphone images from manual microscopes?).

(3) Pay per use (per scan analysed) has a similar low upfront cost to subscriptions. However, the long-term cost depends heavily on the volume of patients scanned. In Africa, the low cost per scan for CAD ($\approx$2 dollars/scan) requires high patient volumes and a pre-screening use case [10]. The choice of use case then drives how the algorithm is designed, trained, and evaluated.

Workforce shortages Africa has a shortage of specialized radiologists and radiographers (Table 1), most of whom work in the private sector and the rest in central hospitals, leaving remote areas with little coverage. In Zimbabwe, more than half of radiologists work in the capital city, the rest in two other cities, and none in rural districts. Thus, it is important to define the location of the use case and understand how the medical imaging product would integrate into the local workflow without dedicated and specialized healthcare professionals. Countries mitigate these shortages in various ways, affecting the potential users of AI tools: Lay cadres in Zimbabwe and Malawi have been trained to use digital X-rays and CAD, and radiographers in Kenya, review the CAD output and send reports, with complex cases sent to a radiologist in referral hospitals for review.

This labor scarcity means that AI can add tremendous value. For example, gestational age and number of fetuses can be accurately assessed with AI from untrained, blind ultrasound sweeps [20].

Supply chains can be hard to change. One malaria diagnostic system above found it useful, from an engineering/device viewpoint, to replace oil immersion with coverslips for slide preparation. But due to the remote settings and difficult supply chains, requiring even a coverslip was a non-trivial barrier to acceptance.

Regulatory AI researchers typically underestimate the importance and complexity of medical regulatory rules and thus omit the necessary groundwork. WHO has published a valuable framework [26], but countries are still assessing

Table 1. Radiologist numbers by country. There is also great intra-country variation.

Country:	Population:	Radiologists:	Radiographers:	Approx population per radiographer:
UK	68,350,000	4,699	45,000	1,500
Zimbabwe	16,340,822	30	450	40,000
Nigeria	227,882,945	400	4,800	50,000
Cote d'Ivoire	31,165,654	13	200	150,000
Malawi	21,104,482	4	30	700,000

the regulation of AI tools. Kenya, South Africa and Uganda regulate these as software as a medical device with no clear framework and risk classification, but rather use ancillary laws like data protection. The EU's CE-mark and In Vitro Diagnostics Regulation (EU-IVDR), and the US FDA approval are currently the most common certifications for CAD software in the African market. But while these may fast-track in-country approval and introduction, local validation is still necessary for the specific location, use case, and imaging equipment. In addition to upfront regulation, it is important to proactively consider the local systems for post-marketing surveillance and quality assurance.

Co-create success metrics that include clinical utility and interpretability This is a crucial (especially for AI researchers) subset of collaborative goal alignment since the AI project needs to track purely medical metrics of success.

1. In one project, hemozoin appeared to be a promising diagnostic biomarker for malaria. But it turned out to be absent in the blood samples of a crucial subset of cases, viz. synchronized *P. falciparum* infections most responsible for deaths [7]. A basic medical performance requirement was unattainable, so despite positive in-lab algorithm results the project was closed.
2. The intended public health impact may extend beyond standard AI notions, and must be clarified upfront. A study of CAD software for tuberculosis (TB) [5] defined these impacts as capacity creation, yield examination, efficiency gain and process optimisation. Identifying these non-AI goals early can inform algorithm and workflow design.

2.3 Data

AI researchers are familiar with how the quality and quantity of data can make or break an algorithm. Indeed, Andrew Ng has argued for a data-centric approach to AI optimization, because careful curation of training data can yield much greater improvements to performance than changes to architectures [19].

Suit the data collection and annotations to the use case The type of data and annotations required depend heavily on the particular use case and location. For example:

1. (Malaria) The diagnosis task requires species identification and thus data from multiple species, whereas drug resistance studies consider only *P. falciparum* [3], but post-drug-treatment malaria parasites are morphologically distinct due to damaged cytoplasm and thus require special post-treatment data. Also, the predominant (and thus clinically relevant) malaria species vary by region, so the required data for AI depends on locale.
2. In a dataset of African cervical images [9], the country-of-origin for many images could be visually identified by regional characteristics or comorbidities [4]. Data variability of this type can greatly impact an algorithm.
3. The expertise of field collaborators can greatly streamline data collection and annotation. For example, a *Loa loa* diagnostic required a dataset rich in very high parasitemia samples (due to the particular use case [12]). The field clinicians in Cameroon were able to efficiently and rapidly collect the required samples, because they knew exactly where the local hot spots were.

Local data effects AI researchers must plan for the targeted region for deployment and account for disparities in data availability and representation, especially in under-resourced regions.

North-East Nigeria, home to ≈30 million people, has only eight functional CT scanners. The border state of Borno is a major medical hub, and patients from Chad and Cameroon often cross the border to access diagnostic services. These cross-border patients have distinct morphological characteristics due to genetic, nutritional, or environmental differences, skewing the dataset. Furthermore, diagnostic infrastructure in nearby states like Gombe and Bauchi is often unreliable, with machines frequently out of service. During such downtimes, patients are redirected to Maiduguri or vice versa, resulting in clusters of imaging data that reflect machine availability rather than true population distribution [1]. If AI models are trained on data collected during scanner downtime periods, the resulting models may reflect biased sampling and may fail to generalize.

Additionally, a supply of local data is required, at a minimum, to test and recalibrate models. For example, inter-clinic variability in slide preparation is perhaps the greatest AI challenge to automated malaria diagnosis [6, 22]. As evident in many of these examples, local data may also be needed for training.

Trust As the saying goes, "Deployment moves at the speed of Trust." A study of stakeholder perspectives [14] highlighted dataset representativeness and contextual equity as key determinants for trust and successful AI for healthcare. An approach to ethical AI development, emphasizing adaptation to local context, is described in [2]. Gaining trust requires a respect for patients' contribution, and often a way to share the benefit of the data. For more details see [11] and [23].

Data is central to AI success, and data collectors, who are usually the field clinicians, deserve more credit. The current model for AI conference papers gives precedence, even in dataset papers, to architectures, and thus AI researchers are typically the first authors. It requires something of a shift in thinking to highlight and reward the vital contributions of clinical partners who collect the data, e.g. by having them as first and anchor authors of dataset papers in AI venues.

2.4 Algorithm design

This is the topic most familiar and dear to our AI hearts. It, of course, requires a great deal of straightforward AI-driven technical expertise. However, it also requires, as an equally necessary condition of success, an expanded set of uniquely healthcare-driven AI tools.

Suit the AI architecture to the local needs An inventory of the exact equipment available in healthcare facilities targeted by the AI project is important, as this can strongly affect architecture and model training decisions. In many African regions, analog X-ray machines have been retrofitted with computed radiography, while machines such as MRI, CT scans, and digital X-rays have limited availability. Most of the up-to-date medical imaging gear is found in the private sector; the public sector has limited functional equipment concentrated in district hospitals, and in rural areas (with over half the population), medical imaging equipment is scarce. Determining which type and brand of medical imaging equipment is available can help ensure integration of the software and minimize the need for countries to procure additional hardware.

Define use case-relevant AI metrics AI is, by nature, quantitative, so the metrics chosen are central to algorithm development, whether for distance metrics, loss functions, or algorithm performance evaluation. We inherit many ready-made, easily applied metrics from the field of AI (e.g. loss functions in PyTorch). But in AI-for-healthcare, metrics must reflect the medical needs to guide the algorithm towards meeting the clinical performance requirements. For example, many AI papers evaluate algorithm performance at the object-level (e.g. parasite thumbnails), whereas clinicians evaluate performance at the patient level. This is a crucial distinction, and object-level metrics (whatever their intermediate value) are not adequate predictors of clinical performance [8, 24]. For a comprehensive (and very readable) inventory of AI metrics for healthcare, see [21].

2.5 Iterative Co-design

A basic key to success is to interact with the local context early and often. Arriving at the clinic door at the project's end with a finished "solution" never works. Some useful techniques include:

1. Conduct feasibility studies with the clinical team for their input into the design and implementation of the product.
2. Conduct facility assessments to better understand the facility infrastructure, patient volumes, and available human resources.

3. Ensure that pilot deployments generate evidence acceptable to the medical world, not just to the AI world [25].
4. Work with MoHs or international programs for the specific disease areas, to get their buy-in and to understand the local context of deployment/use.
5. Keep clinicians looped in, for example, in full-team sync meetings, even when not in the field. A clinician may see elephants in pink tutus when the AI team sees nothing special – and vice versa. This is a crucial way to head off pitfalls and, conversely, to leverage opportunities.

It is desirable to get into the field/clinic early and often to test prototypes, because important (even existential) reality checks always arise. For AI field tests, two common findings are (a) out-of-distribution data, and (b) unforeseen conditions in algorithm use. The earlier these situations are detected and incorporated into development, the better. Field tests always yield new, high-value information. For example:

1. In the *Loa loa* project described above, a field trial surfaced the issue that blood samples sometimes coagulated before being imaged, giving undercounts that would be disastrous in the targeted use case. This revealed the need for a built-in coagulation detector module.
2. A malaria project using Acridine Orange stain was highly effective, from a purely in-lab algorithm perspective. But during the first field test, the clinical teams made clear that AO stain was simply impractical for field use, an issue that led to shuttering the project.

3 OK, but let's be realistic

Addressing all the items listed above is ideal, but requires a team with substantial local connections, which most of us do not have. However, certain concrete actions, accessible to almost all researchers, can ensure that an AI-for-healthcare project is well-informed by local conditions and domain expertise:

1. At career transition points (e.g. choosing a PhD program), seek out labs and teams that have clinical connections: labs with adjoined research medical schools, with clinicians on the team, or with a track record of field work or advanced-stage projects.
2. Choose an AI problem based on your existing connections to local contexts: That is, let the AI project flow out of the needs of a local situation. This helps ensure that the problem being addressed is real, and that domain experts are in the loop from the start.
3. Subordinate algorithm design to the medical specifics. The best algorithmic solution might use off-the-shelf methods, in which case the novelty lies in its real-world utility.
4. Consult the *non-AI* medical literature - this is essential to understanding the AI task and is readily available. Sources include the medical literature on particular illnesses, and literature on health systems (e.g. local NGOs, PATH, CHAI, PLOS, ASTMH). Also, cite this literature in your AI paper to benefit other AI researchers.

5. Attend non-AI conferences, or those workshops at AI conferences with clinical presence, to meet collaborators, learn medical use cases, and the urgent problems AI can solve. E.g.: CLINICCAI at MICCAI; UnionConf (TB); ICASA (HIV); ASTMH (tropical medicine); IASLC Conference (lung cancer).
6. Define clinically relevant metrics for use in AI development for the targeted use case (an excellent resource for choosing appropriate metrics is [16]). This may require new construction or mathematical derivation, since these metrics often do not exist for a given medical use case in AI-usable form (for a concrete example from malaria, see [8]).

AI lifecycle mapping of our framework: In problem definition, engage local experts to co-create clinically relevant goals (e.g., Zambia cervical cancer shaped by HIV); in data and model development, curate context-specific datasets, use patient-level metrics, and design for local hardware (e.g., CT downtime bias in North-East Nigeria); and in evaluation and deployment, validate with clinical standards, pilot through field testing, and adapt to pricing and regulatory systems (e.g., *Loa loa* trial revealing coagulation detector needs).

AI expertise is a powerful and highly valued skill set. AI researchers can be welcome and high-value members of medical teams that accelerate progress in healthcare, if we carefully attend to the medical context.

Acknowledgments. Partially funded by Global Health Labs, Inc. (www.ghlabs.org).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Adejoh, T., Onwujekwe, E., Chiegwu, H., et al.: Computed tomography scanner census and adult head dose in nigeria. The Egyptian Journal of Radiology and Nuclear Medicine (2018)
2. Amugongo, L., Kriebitz, A., Boch, A., Lütge, C.: Operationalising AI ethics through the agile software development lifecycle: a case study of AI-enabled mobile health applications. AI and Ethics (2025). https://doi.org/10.1007/s43681-023-00331-3
3. Ashley, E., , Mehul Dhorda, M., White, N., et al.: Spread of artemisinin resistance in *Plasmodium falciparum* malaria. New England Journal of Medicine (2014). https://doi.org/10.1056/NEJMoa1314981
4. Asiedu, D.M.: Personal communication (2024)
5. Creswell, J., Vo, L., Qin, Z., et al.: Early user perspectives on using computer-aided detection software for interpreting chest x-ray images to enhance access and quality of care for persons with tuberculosis. BMC Global Public Health (2023)
6. Das, D., Vongpromed, R., Dhorda, M., et al.: Field evaluation of the diagnostic performance of EasyScan GO: a digital malaria microscopy device based on machine-learning. Malaria J (2022)

7. Delahunt, C., Horning, M., Wilson, B., et al.: Limitations of haemozoin-based diagnosis of *Plasmodium falciparum* using dark-field microscopy. Malar J. (2014)
8. Delahunt, C., Gachuhi, N., Horning, M.: Metrics to guide development of machine learning algorithms for malaria diagnosis. Frontiers Malaria (2024)
9. Dotson, M.E., Asiedu, M., Ramanujam, N.: Speculum-free callascope for cervical self-visualization: Acceptability, feasibility, and improved awareness of the reproductive system. JCO Global Oncology (2020)
10. Du, Y., Greuter, M., Prokop, M., de Bock, G.: Pricing and cost-saving potential for deep-learning computer-aided lung nodule detection software in CT lung cancer screening. Insights Imaging (2023)
11. IDDO: https://www.iddo.org/governance/data-access-committee, accessed 25 June 2025
12. Kamgno, J., Pion, S., Boussinesq, M., et al.: A test-and-not-treat strategy for onchocerciasis in *Loa loa*-endemic areas. New England Journal of Medicine (2017). https://doi.org/10.1056/NEJMoa1705026
13. Koller, D., Bengio, Y.: A fireside chat with Daphne Koller at ICLR (2018), https://www.youtube.com/watch?v=N4mdV1CIpvI
14. Kuo, R., Freethy, A., Furniss, D., et al.: Stakeholder perspectives towards diagnostic artificial intelligence: a co-produced qualitative evidence synthesis. eClinical Medicine (2024)
15. Lekadir, K., Frangi, A., Porras, A., et al.: Future-ai: international consensus guideline for trustworthy and deployable artificial intelligence in healthcare. BMJ (2025). https://doi.org/10.1136/bmj-2024-081554
16. Maier-Hein, L., Reinke, A., et al.: Metrics reloaded: Pitfalls and recommendations for image analysis validation. arXiv (2022), https://arxiv.org/abs/2206.01653
17. Mensah, Y., Nkyekyer, K., Mensah, K.: The Ghanaian woman's experience and perception of ultrasound use in antenatal care. Ghana Medical J (2014)
18. Mwanahamuntu, M., Parham, G., et al.: Advancing cervical cancer prevention initiatives in resource-constrained settings: insights from the Cervical Cancer Prevention Program in Zambia. PLoS Med. (2011)
19. Ng, A.: Practical limitations of todays' deep learning in healthcare. ML4H conference (2020), https://nips.cc/virtual/2020/19127
20. Pokaprakarn, T., Stringer, J., et al.: Ai estimation of gestational age from blind ultrasound sweeps in low-resource settings. NEJM Evidence (2022)
21. Reinke, A., Tizabi, M., et al.: Understanding metric-related pitfalls in image analysis validation. Nature Methods (2024)
22. Torres, K., Bachman, C., Gamboa, V., et al.: Automated microscopy for routine malaria diagnosis: a field comparison on Giemsa-stained blood films in Peru. Malaria J (2018)
23. Understanding Patient Data: https://understandingpatientdata.org.uk/, accessed 25 June 2025
24. Varoquaux, G., Cheplygina, V.: Machine learning for medical imaging: methodological failures and recommendations for the future. NPJ Digital Medicine (2022), https://doi.org/10.1038/s41746-022-00592-y
25. WHO: Generating evidence for artificial intelligence-based medical devices: a framework for training, validation and evaluation (2021), World Health Organization, Geneva, Switzerland
26. WHO: Regulatory considerations on artificial intelligence for health. World Health Organization (2023)

27. Wikipedia: Article on PCPND https://en.wikipedia.org/wiki/Pre-Conception_and_Pre-Natal_Diagnostic_Techniques_Act,_1994, accessed 25 June, 2025