

Artificial intelligence and health equity in primary care: A qualitative study with key stakeholders

Alexander d'Elia ^{a*}, Mark Gabbay ^a, Lucy Frith ^b, Sarah Rodgers ^a and Ciara Kierans ^a

^a *Institute of Population Health, University of Liverpool, Liverpool, United Kingdom;*

^b *Centre for Social Ethics & Policy, University of Manchester, Manchester, United Kingdom*

*Corresponding author:

Dr Alexander d'Elia:

adelia@liverpool.ac.uk

3rd Floor, Whelan Building

University of Liverpool

Brownlow Hill

Liverpool

L69 3GB

United Kingdom

Artificial Intelligence (AI)-augmented interventions are currently being rolled out across primary care, but how it affects health equity remains insufficiently understood. This qualitative study addresses this gap through an ethnographical inquiry based on 32 interviews and focus groups with stakeholders including commissioners, decision makers, AI developers, researchers, GPs and patient groups involved in the implementation of AI in English primary care. We took a sociotechnical perspective in order to assess how the stakeholders can improve health equity through the implementation process of AI within the wider system. We found that regulation and policy alone cannot guarantee equitable implementation of AI but can provide a framework to enable other stakeholders to take measures to promote equity: fostering a shared understanding of the causal mechanisms of AI and health equity, how to measure health equity, and how to share data necessary for equity promotion. Further, all stakeholders need to be on board for equitable implementation, and currently innovation leaves clinicians and patients behind. Capacity building is needed to achieve this, in particular at local commissioning and clinician level. Careful implementation and pragmatically focused research are needed to make AI in primary care capable of advancing health equity.

Keywords: Artificial Intelligence, Implementation, Health Equity, Primary Care

Wordcount: 6928 (incl. abstract, tables, references and captions, excl. appendices)

Introduction

Artificial intelligence (AI) can be described as a computer system that performs tasks traditionally requiring human intelligence. Everyday examples include predicting preferences in social media feeds, recognizing faces in photos, or recently, compiling complex information in natural language such as ChatGPT¹. A rapidly expanding field, AI-augmented interventions are considered integral to the healthcare of the future and current applications include interpreting X-rays and ECGs. Widespread application of AI-based innovations² in healthcare is already happening, and primary care is no exception.

At the same time, health inequity (hereafter *inequity*) is increasingly discussed, not least in the context of the ongoing COVID-19 pandemic³. Through potentially freeing up resources and enabling more personalized care, AI has been described as a potential enabler for more equitable healthcare². However, as AI interacts with socioeconomic, gender and ethnic health inequities on many different levels, it could both increase or decrease inequities, depending on application and implementation^{4,5}. Primary care holds a unique role in tackling inequity. It can be both a source and a magnifier of inequities, as well as a platform for mitigation⁶. For the purpose of this study, primary care is defined as primary care services provided to individual patients, specifically in the English National Healthcare System (NHS), and not including wider public health policy⁷.

In the context of inequity and implementation, AI differs from other digital interventions in a number of ways. The most important of these are: in contrast to other complex health-informatics interventions; AI-augmented interventions are likely to be more opaque in showing how outputs are achieved⁸; AI has the potential (largely untapped at the moment) to be self-improving, raising issues around evaluation⁹; AI opens up new domains of healthcare provision to digitalization and privatization, which

may have unforeseen effects on the primary-care system and on health equity^{10 11}.

Finally, the pace of innovation brought on by AI is unprecedented in healthcare, further elevating long-standing concerns about technology implementation in healthcare (e.g. poor adaptation and adverse system effects¹²)¹³.

This study is underpinned by a comprehensive literature review describing the pathways for how AI can affect inequity in primary care¹⁴, producing a framework for how AI may affect inequity *in and through* primary care (Appendix 1). Our review found a multitude of potential mechanisms, broadly divided into ‘intrinsic’ and ‘extrinsic’ effects. Intrinsic effects include inequity impact from skewed, or biased, data collection and algorithms that may contribute to the different quality of care different population groups might receive (e.g., Obermeyer et al¹⁵). Extrinsic effects on the other hand take place through the implementation of a given AI-augmented intervention and the interaction between the AI-intervention and the system it is deployed in. A typical example is the ‘digital divide’; the tendency for socioeconomically disadvantaged groups and the elderly to have less access to technology and thus benefitting less from AI interventions requiring such technology compared to their more advantaged peers. Our review also discussed potential extrinsic inequity issues stemming from a lack of patient trust, increased focus on self-management (the success of which correlates to socioeconomic), adding workload to already strained clinical workforce causing out-crowding effects, and a general risk that interventions target uncomplicated, easy patients over patients with complex socially multimorbidity. If these inequity-exaggerating effects are to be avoided or mitigated, and conversely if the potentially equity-positive effects are to be gained, careful implementation of AI products is necessary, which in turn requires a thorough understanding of the system in which they are being deployed. The importance of taking this social context into account when

implementing health technology has been extensively studied. For example a 2017 meta-analysis found that sociotechnical factors were involved in more than two thirds of health technology interventions that failed to achieve their objectives^{12 16}.

There are numerous reports on how to best implement AI in healthcare and other sectors, typically by governmental and non-governmental organisations^{2 17-22}. Being typically general and (intentionally) speculative they provide a starting point for further research, but few provide empirically grounded findings. Among the more ambitious of such projects is a recent report by The Ada Lovelace Institute, ‘The Knotted Pipeline: Data-driven systems and inequities in health and social care’²³. The report builds on interviews with organizational stakeholders in the health-data context (including but not exclusively AI) and concludes that there is a distinct need for a common definition of equity and for improved collaboration between different organisations in order to improve health equity outcomes thorough health data. The report also emphasizes the perils of overestimating the ability of data-driven technologies in improving outcomes when they are being implemented in an existing, highly complex system; the risk of viewing AI as a panacea and failing to take the wider organisational and social environment into account. In contrast to the Ada Lovelace piece, we focus on AI and primary care specifically, apply a different and specific research approach for data collection and analysis, and involve a diverse set of stakeholders, including patients. To summarize, our study aimed to assess how the sociotechnical network in which UK primary-care AI-implementation is taking place (i.e., the organizations, regulations and stakeholders involved) is conducive to making AI a force for health equity, and what makes for a favourable environment for reducing health inequity.

Materials and Methods

This is a qualitative, explorative, social-constructivist study drawing primarily from 27 individual interviews and five focus groups with stakeholders involved in the implementation of AI in primary care.

Participants

The selection of participants was both purposive and theory driven. The purposive sampling meant a broad representation of different stakeholders, whilst the theory-driven aspect allowed for continuous recruitment as the study went on with interviewees being recruited to explore emerging themes. In practice this meant deciding early on which broad groups were seen as integral stakeholders and thus important to include (the purposive aspect), but also to draw on the findings from the interviews and recruit sequentially to explore the themes that were developing (the theory-driven aspect). For example, early interviews showed that algorithmic impact assessment (AIA) were a key component, and thus people with specific experience of this were recruited (e.g., #17 from the Ada Lovelace Institute). Participants consisted of:

- *Developer* stakeholders. Product developers working on primary-care AI products. Recruited by contacting a range of UK AI companies involved in primary care via public email addresses.
- *General Practitioners* (GPs) Recruited by disseminating the invitation through South Birmingham GP Training Scheme mailing list, and through snowballing from interviewees.
- *Commissioning* stakeholders focusing on Integrated Care System (ICS) and Academic Health Service Networks (AHSN) members involved in digital

product procurement. Recruited through emailing all ICSs and AHSNs across England, of which a minority were available for interviews.

- *Regulator* stakeholders representing the organizations central in implementation process, including the Medicines and Healthcare products Regulatory Agency (MHRA), the National Institute of Care Excellence (NICE), the Department of Health and Social Care (DHSC), the Care Quality Commission (CQC) and the Information Commissioners Office (ICO), as well as the advisory non-governmental Ada Lovelace Institute. Recruited through emailing respective organization.
- *Researcher* stakeholders (i.e., researchers investigating topics of AI and its impact on healthcare), including both clinical academics as well as non-medical researchers. Recruited through targeted emails and subsequent snowballing.
- *People living with diabetes mellitus* (insulin- and non-insulin dependent) (PwDs). Recruited via Diabetes UK's West Midlands office, who in turn asked local organizations to help with the recruitment of participants for focus group sessions during their local meetings around the West Midlands, England. PwDs were recruited as a case study for patients with chronic illnesses in general, and the emphasis was on the participatory component of their own healthcare and their relationship with primary care and AI.

To capture divergent ideas and reach adequate fidelity for every interviewed group, it was decided that every stakeholder group would have at least five interviewees. However, only four commissioning stakeholders were interviewed, due to a last-minute cancellation.

Inclusion and exclusion criteria

Participants from the groups representing their professions/organisations were not selected based on demographic characteristics. No individual demographic selection of PwDs included was made, however PwD groups were purposively sampled to an extent as the research question necessitated diverse representation (a minimum of 1/3 of either gender, and 1/5 of minority ethnicity for all PwDs). Exclusion criteria for all participants included:

- Lacking capacity to consent to participation.
- Not proficient in English (due to pandemic-related limited resources of clinically confidential interpretation services it was not feasible to offer this).

Methodology

This study applied an iterative constant-comparative methodology, as described by Silverman²⁴, allowing for an effective method for conducting explorative research from a series of sequential interviews²⁵. Analysis was conducted using three levels of coding in an iterative manner, inspired by Grounded Theory²⁶; herein referred to as codes, themes and theory (technically a middle-range theory).

Initially, the implementation systems and processes were mapped in general terms using publicly available documents and complemented by interview data as well as a conceptual model of equitable implementation (appendix 2 and 3, respectively), to provide a backdrop for the rest of the study. The collection and analysis of the interview data thereafter consisted of six steps, repeated until saturation (Figure 1). All steps were conducted by the first author with advisory support from the co-authors, apart from step 3:

- (1) Data collection: interviews and focus groups. All data collection sessions were conducted online using Zoom videocall software ²⁷, with the exception of three of the PwD focus groups, which were conducted at local Diabetes UK meetings. Diabetes UK sessions were not recorded, but field notes were taken. Both interviews and focus group sessions were semi-structured, with different interview schedules depending on what themes were being explored. Transcription was conducted using Sonix software ²⁸ with subsequent manual editing. Anonymization took place after transcription with each transcript given a code (Table 1) and recording deleted. Data was then imported into NVivo 12.0 qualitative analysis software for subsequent steps ²⁹.
- (2) Familiarization; a read through of the data without coding. Comments were made based on free association.
- (3) Initial, free coding, of transcripts followed by comparison with all the previous codes to decide if a new code should be created or the excerpt should be coded under an existing code. The coding was predominantly conducted by the first author, but two transcripts were double coded by CK and MG to improve internal validity.
- (4) Codes were then collated into themes. As with the initial codes, all new themes were compared to the previous themes avoid duplication, all existing themes were reviewed in each round, and altered if needed to remain compatible with all initial codes. (See appendix 4 for examples of iterations of the theme list).
- (5) Building provisional middle-range theory from themes; the theory acted as a glue that related the themes to each other, keeping the inquiry focused on the research question. The theory was revised continuously in line with the development of the themes.

- (6) Recruit new participants to explore the themes and theory further; develop interview guides on new lines of inquiry. Again, this recruitment was partially guided by the theory, partially by the criteria of including all stakeholder groups.

These steps were repeated as new data were collected, until saturation.

Finally, the results were compiled as a narrative, building on the final iteration of the themes and the joining theory, and compared to preceding scholarship. A graphical logic model serves to visualize the theory (Figure 2), which has been used in studies with similar methodology³⁰⁻³².

Public involvement

Patient and/or public involvement (PPI) in research and development and implementation of medical AI has been suggested as a way of improving trust and reducing the risk of adverse health-equity effects^{15 33}, as has been previously discussed in non-AI contexts^{34 35}. In the context of medical AI, previous research has shown a link between PPI to improve public understanding of AI and acceptability of a hypothetical AI product, predicting that this would lead to more successful implementation³⁶.

During the development and carrying out of this study, two remunerated public advisors were involved, helping to shape the direction of the study, the interpretation of the results (i.e., the Discussion section) and the dissemination. The advisors identify with traditionally marginalized populations (one of British Asian ethnicity and one registered disabled and member of the LGBT community) and helped widen the perspectives of the inquiry. On a larger scale, the two advisors have been involved in the planning of the PhD project this study forms part of, providing input on useful areas of inquiry, overarching methodology and dissemination.

Reflexivity

The first author is a general practitioner in training in the English NHS. Subsequently, the study has been conducted from the perspective of a practical primary care practitioner. It is likely that the findings would have been slightly different if a commissioning stakeholder or private AI company had conducted a similar study, due to differences of perspective. Second, the main author is a White, 34-year-old, healthy, medically educated male; a privileged group in terms of inequity. This clearly affects understanding of inequity, even as the same has considerable clinical and academic experience of inequity. The two public advisors' background helped broaden perspectives.

Ethics

This study was approved by The University of Liverpool, research ethics committee, reference number 10229. The study was conducted in accordance with the recommendations for physicians involved in research on human subjects adopted by the 18th World Medical Assembly, Helsinki 1964, and later revisions.

Results and Interpretation

Demographics

Twenty-seven individual interviews and five focus group sessions (all PwD sessions plus one of the commercial developer sessions, #20) were conducted during ten months (April 2022 to January 2023). Of these, GPs were the largest group, with seven interviews conducted, followed by researchers (n=6), regulatory and advisory stakeholders (n=6), commercial developers (n=4 plus 1 focus group), local commissioners (n=4) and PwD focus groups (n=4). All these sessions were conducted remotely. Demographic information was not collected for the professional stakeholders

(i.e., all but PwDs) beyond gender, which ranged from 17 % female (researchers) to 57 % female (GPs). For all professional stakeholders 36 % were female. The average duration of the interviews was 39 minutes (ranging from 23 to 63 minutes).

The four PwD groups ranged in size from six to ten participants with a total of 29. All but one of these sessions (PwD 4) were conducted in person at the respective Diabetes UK meeting location. For PwDs basic information on age and ethnicity was collected to meet the diversity criteria of the study. Of all PwD participants, 55 % were female (33-67 %), 86 % were over 65 years old (70-100 %) and 21 % were of non-White-British ethnicity³⁷ (0-50 %). The average focus group lasted for 43 minutes (33 to 51 minutes).

Addressing the aim

Our aim is to assess how the sociotechnical network in which UK primary-care AI-implementation is taking place (i.e., the organizations, regulations and stakeholders involved) is conducive to making AI a force for health equity.

Comparing the data and derived themes and joining theory (Figure 2) with the mentioned model of equitable implementation (appendix 3), a complicated picture of the current sociotechnical network emerges. Many of the actions indicated as required in this ‘ideal’ model indeed appear to take place; algorithmic impact assessments are developed to assess the impact of AI in various settings (although no primary care specific ones exist), initiatives to improve information flow between different stakeholders are taking place (e.g. the AI and Digital Regulations Service/ADRS and AHSNs). Arguably, looking beyond AI, this push for local involvement is one of the reasons for the recent reorganization of local NHS groups into larger regional Integrated Care Systems.

Our interview and focus group data showed that more needs to be done for the system to be conducive to equitable implementation. The five themes discussed in this section

represent five interconnected approaches to ‘catalyze’ such an implementation chain, and the above graphical logic model (Figure 2) – the *theory* – illustrates how they relate to each other.

Theme 1: Equity through influence

Starting from the left side of the theory model (Figure 2), *equity through influence* encompasses the inherent power imbalance between the different stakeholders when shaping a new intervention, and AI interventions in particular. Extensively studied and described in other healthcare contexts³⁸⁻⁴⁰, equitable implementation of technology needs to be addressed across the hierarchy of stakeholders, simultaneously. Applying an implementation science perspective⁴¹ leads to the conclusion that in order for an intervention to function well (and in the context of the NHS this includes promoting equity), an intervention must be anchored in the behavior of all its stakeholders, and the communication culture between these stakeholders. As one interviewee involved in an advisory organization put it:

I think there needs to be a much more sophisticated public dialogue around applications of AI and health, the public need to be joined in with it. The government needs to be participating in this conversation, and it shouldn't be consultative [i.e., not just a one-time exercise]. We shouldn't just throw out something and say, this is happening, deal with it. There needs to be sustained, meaningful public engagement in this space. I think that is the number one takeaway from my work.

#17, Ada Lovelace Institute

Consequently, ‘bringing all stakeholders on-board’ and making sure that the intervention responds to all stakeholders’ needs throughout both development, implementation and evaluation is essential. Involving the public (patient and public involvement/PPI) is needed to understand local needs and local inequity:

I think a lot of it will come down to the creators [of AI products] to have a greater understanding of the needs of the population that they're aiming for. And that's how you can make things and bespoke to the context. I think some of these companies must actually go into primary care to see what happens at the reception to kind of see what the queries are from patients and stuff.

#1, GP

A PwD expressed frustration over how new interventions in primary care did not seem to be in the interests of the patients:

All that is needed is some common sense. If they knew how we experienced [living with diabetes] they would be able to [provide care] better!

PwD, group 1

However, both in the context studied in this project and in general, there is a lack of consensus on how to best conduct PPI for improved equity in outcomes⁴², which needs to be addressed. Participatory primary care clinics, such as the Good Things Foundation's NHS Digital Health Labs⁴³ offer an interesting approach. This initiative, which ran between 2017 and 2020 in the South of England, included 'Digital Health Hubs'; open spaces where community members could interact with the care system (e.g., their primary care practice) with various digital tools. As part of this they were studied and interviewed, both in the moment and as structured focus groups, and the mainly focusing on accessibility of digital user interfaces. Similar set-ups could be used to locally evaluate AI systems for 'social compatibility', i.e., accessibility, acceptability, and ability to meet local needs. Such an approach would need to ensure a representative sample: There is an inherent risk of bias when recruiting mainly the 'digitally interested', potentially worsening inequity.

Theme 2: Shared understanding

A key part of *equity through influence* is *shared understanding*; to achieve more equitable influence over the implementation process, there needs to be a consistent understanding of core concepts, such as ‘*what is health (in-)equity?*’ and ‘*what is the goal of the given intervention?*’. Again, this mirrors implementation science frameworks such as normalization process theory⁴⁴. There were differing thoughts on the role of AI in primary care among the participants; whilst AI developers seem to consider their products as largely ‘agnostic’, socially passive tools (as reflected in the focus on increased efficiency/productivity), they are being implemented as part of a complex system with multiple sometimes conflicting goals, as emphasized by the academic interviewees among others. An interviewee at an AI company echoed this view of their product as socially inert, and that affecting inequity was beyond their reach:

We have worked on how on how to make the system easier for patients with poor vision, for example, but we have not focused on the broader structural aspects, such as equality in-between women and men in usage. I think this is something that we all as doctors have in the back of our heads, however.

#20, AI company 1

A parallel can be drawn to Han et al’s⁴⁵ ethnography on a critical care IT system which was found to be increasing mortality due to a discrepancy between the perceived usage conditions (by the developers and implementers) and the actual situation (namely a reliance on informal communication and routines that were not replicated in the IT system), or Greenhalgh et al’s³⁸ rapid ethnography on the failed implementation of home-based health-IT for geriatric patients. In both cases, implantation failure arose from a discrepancy in how the producing and implementing stakeholders assumed that the products would be used and for which goals, and the reality of the users’ context. In

some cases, the concept of health equity (equality in health outcomes) is in conflict with outright utilitarianism (the greatest good for the most patients), which can create tensions in designing and implementing interventions:

A lot of health is built off a utilitarian view: the greatest good for the greatest number.

If you're following that, if you're truly following that, then you'd say, well, actually, if it caused minor harm to a few people, if it caused greater good for a lot of people, then you should probably go with it.

#9, GP and AI researcher

Clearly then, there is a link to the previously discussed theme *equity through influence*; in order to design and implement an AI system concordant with the NHS goal of reducing inequity in and through primary care, all stakeholders need to speak the same language, sharing experiences and creating a shared understanding of the context, the goals and the intervention.

Theme 3: Shared frameworks and collaborative platforms

In order to develop such a *shared understanding*, there needs to be standards and platforms in place: *shared frameworks*. This refers to both ways and regulations around sharing data as well as organizational structures to bring in local perspectives and feedback into commissioning, production, and regulation. Algorithmic impact assessments provide tools for a standardized and relatively convenient assessment of AI-products by commissioning and evaluating stakeholders. Current examples are rare but exist, e.g., one is being trialed by Ada Lovelace Institute at the moment ⁴⁶. For equity improvement, future developments need a stronger and more detailed focus on equity and could draw from equity assessment tools such as the Health Equity Assessment Toolkit (HEAT ⁴⁷):

It's almost irrelevant whether it's a public sector or private sector or anything else:

Using the HEAT is very much about understanding your audience, understanding the outcomes you want to achieve, and utilizing these principles by working with different professionals to get to the core of our audience and make a difference. That's how I view it.

#4, DHSC

Such AIAs could both act as a foundation for a common understanding and a practical guide for assessing AI products, before and after implementation. To develop such a tool, and to enable a *shared understanding* of how to address inequity, collaborative structures are needed. The AI and Digital Regulations Service (ADRS) is an example of a high-level collaborative system, while AHSNs provide potential starting points for more local entities. An improved data sharing framework is needed to enable said assessments and facilitate more equitable AI in the future (through access to real-life data on local populations), created in collaboration between regulatory organizations but presumably led by the ICO whilst maintaining privacy and subsequently trust:

Data access is a big issue. The Goldacre review [A 2022 UK government-commissioned health-data whitepaper] advises to go close to the data source and do the research and development close to the implementation e.g., build systems at GP level. That enables easy trialing through running a system in parallel to real GP.

#10, MHRA

Theme 4: Capacity building and knowledge generation

To accomplish the above and enable equitable implementation, *capacity building and knowledge generation* is needed across the whole ecosystem. This means different things for different stakeholders but highlights the need for resourcing and a coherent push for standards and understanding (e.g. how to assess an AI system from an equity perspective) and how to best conduct PPI for equitable AI, including increased capacity

for assessment of new technology (e.g. MHRA), national evaluation (e.g. NICE), local assessment and evaluation (e.g. ICSs, primary care centers and networks) and through academia. An interviewed GP commented on the outlook for involving her practice in AI implementation:

Am I super enthusiastic for that at the moment? No, because we don't have the capacity. If as a cluster, we had funding for it and there was proper time to resource it, then yes, because you know, there's a real chance that it might make a meaningful difference to patients and practices.

#11, GP

In the primary care context, local ‘champions’ – highly motivated individuals or local organisations driving change on their own initiative – was mentioned as a way of achieving a collaborative and equitable implementation. However, such developments may also worsen inequity:

There are the big centers where they've got a lot of money and a lot of expertise and there are places that people want to work. And basically, those are the places that are probably developing their technologies and adopting AI and being sort of mooted as the AI leaders across healthcare and the smaller trusts that struggle to recruit, that probably serve more deprived populations, have more challenges about providing healthcare and access to healthcare, and don't have the money all the time or the expertise in order to develop these technologies or adopt and deploy them. I think they will just fall further and further behind.

#14, clinician and AI researcher

This again emphasises the need for a high-level framework for AI implementation to promote equitable opportunities across localities with different circumstances.

Theme 5: Division of responsibility

Finally, the *division of responsibility* is discussed, depicted on the right side of the theory logic model (Figure 2). In the model, the stakeholders are presented as an organizational hierarchy of a typical AI intervention, from production to use. Different approaches to equity promotion are needed at different levels of the ecosystem, but (as depicted in the figure with a large double-sided arrow connecting the stakeholders with *equity through influence*), all stakeholders need to feed into a common understanding and approach. The one-sided arrow in the model depicts how the (inter-) national policy and regulation environment have a particular role to play in enabling *equity through influence, shared understanding, and shared frameworks*, as well as *capacity building and knowledge generation*.

In other words, high-level political and regulatory oversight is needed to accomplish the other themes, not by top-down action but through providing a clear set of regulations and protocols to enable a synchronized approach to be conducted at intermediate and local levels, in close collaboration between users, commissioners and producers. As described by an interviewee at MHRA:

Ultimately it is very difficult for regulations to ensure equity. Local expertise is needed to assess it in the context [a product] is implemented. On the same note, MHRA has no agency to perform post-market surveillance; it is up to the user to follow up.

#10, MHRA

Meanwhile, local stakeholders such as ICSs, primary care centers (and -networks) and AHSNs need to lead local implementation and involvement to produce interventions suited to the local needs, and subsequent evaluation; *applying and utilizing* the frameworks built by the higher system levels. Several interviewees talked about ‘place based’ implementation (as per NHS England an area of about 30 000 inhabitants⁴⁸), typically consisting of a large borough or a primary care network:

I think practices are too small and the majority of ICSs I think are too big. What I want to do in North [City] is very different because that population has very little in common with East [City] population, for example. There are very few ICSs where they've got common issues across the whole economy; they tend to be in community-based sort of pockets. So, I think it has to be at a 'place' level.

#6, ICS digital strategist

As mentioned, this is in line with the general ambition for ICSs to act as an umbrella and supporting organization for localized initiatives.

Discussion

This paper brings together a set of organizational and public stakeholders in an ethnographically driven, explorative study. This approach, combined with the iterative constant-comparative methodology, allowed us to derive a theory from a broad range of empirical observations and create a uniquely holistic foundation for improvement of the AI implementation process for primary care, as set out in our key messages below. The study has several limitations: An inquiry assessing the intersection of AI, inequity and primary care will inevitably remain superficial to a degree, and this holistic system perspective is both a strength and weakness. Further, the study lacked demographic sampling for diverse participants among GPs and PwD groups. However, a 'minimum' demographic diversity criterion was applied to avoid an overly homogenous sample of PwD groups.

Through analysing the primary data against the frameworks outlined in the method section, a few key messages can be identified:

- The organizational ecosystem is largely unprepared for equitable implementation of AI in primary care. However, as mentioned the situation is multifaceted; strong systems for PPI, AIAs and a general movement towards

localization of commissioning through ICSs and AHSNs is promising, but what is lacking is a unifying strategy encompassing the themes discussed above.

- High-level (inter-)national policy and regulation must promote frameworks for equitable implementation that enables all stakeholders, from developers to GPs and patients, to conduct implementation in an equitable manner. Such frameworks enable AIAs and implementation processes to be conducted at a local level, in turn enabling contextualization and local stakeholder involvement.
- Local capacity to conduct such processes subsequently need strengthening, and further research is needed to define how such AIAs and frameworks are best designed, including into means to effectively quantify and evaluate inequity effects of AI, and how to best involve the public in shaping AI interventions for health equity.

Whilst the difficulty of implementing technological interventions in complex systems is previously well studied^{13 41}, the context of AI, inequity and primary care brings further complications. More than anything, the study highlighted the complexity of implementing AI in primary care, and the difficulties of assessing such processes from an inequity perspective. Part of the complexity stems from the relative absence of delineation; the term ‘AI in primary care’ as used here is very wide (as outlined in the introduction section). More pragmatically aimed future research will naturally have to narrow the definition to certain AI interventions or well-defined groups, to clearly be able to lay out the *mechanism of change* of a given intervention and the potential equity effects.

AI interventions in patient-facing healthcare will remain highly complex and will need to be studied in a manner that takes the sociotechnical complexity into

account. Sociotechnical approaches such as ethnographically inspired inquiry certainly have an important role to play in future explorations.

Acknowledgement

We acknowledge the valuable contributions by the two public advisors, Irum Durrani and Adele Thomas, as outlined above.

Funding

This study was conducted as part of the PhD project “Artificial Intelligence and Health Inequities in Primary Care”, by Alexander d'Elia. The PhD project is funded by Applied Research Collaboration Northwest Coast (ARC NWC), in turn funded by the UK National Institute for Health Research (NIHR).

Declaration of interests

The funder or the authors have no conflicts of interest to declare.

References

1. Stuart R, Peter N. Artificial Intelligence: A Modern Approach. 2020: Prentice Hall, Upper Saddle River, NJ.
2. Joshi I, Morley J. Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care. NHSX, 2019.
3. Bambra C, Riordan R, Ford J, et al. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health* 2020;74(11):964-68. doi: 10.1136/jech-2020-214401 [published Online First: 2020/06/15]
4. Veinot TC, Mitchell H, Ancker JS. Good intentions are not enough: how informatics interventions can worsen inequality. *J Am Med Inform Assoc* 2018;25(8):1080-88. doi: 10.1093/jamia/ocy052 [published Online First: 2018/05/23]
5. Academy of Royal Medical Colleges. Artificial Intelligence in Healthcare: Academy of Royal Medical Colleges, 2018.
6. Popay J, Kowarzik U, Mallinson S, et al. Social problems, primary care and pathways to help and support: addressing health inequalities at the individual level. Part I: the GP perspective. *Journal of Epidemiology & Community Health* 2007;61(11):966-71.
7. Muldoon LK, Hogg WE, Levitt M. Primary care (PC) and primary health care (PHC). *Canadian Journal of Public Health* 2006;97(5):409-11.
8. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence* 2019;1(5):206-15.
9. Joshi I, Morley, J. Artificial Intelligence: How to get it right. Putting policy into practice for safe data-driven innovation in health and care.: NHSX, 2019.
10. Arora A. Conceptualising artificial intelligence as a digital healthcare innovation: an introductory review. *Medical Devices (Auckland, NZ)* 2020;13:223.
11. McCartney M. General practice can't just exclude sick people. *BMJ: British Medical Journal* 2017;359
12. Kim MO, Coiera E, Magrabi F. Problems with health information technology and their effects on care delivery and patient outcomes: a systematic review. *Journal of the American Medical Informatics Association* 2017;24(2):246-50. doi: 10.1093/jamia/ocw154

13. Coiera E. The Price of Artificial Intelligence. *Yearb Med Inform* 2019;28(1):14-15. doi: 10.1055/s-0039-1677892 [published Online First: 2019/04/26]
14. d'Elia A, Gabbay M, Rodgers S, et al. Artificial intelligence and health inequities in primary care: a systematic scoping review and framework. *Family Medicine and Community Health* 2022;10(Suppl 1):e001670.
15. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366(6464):447-53. doi: 10.1126/science.aax2342 [published Online First: 2019/10/28]
16. Shaw J, Rudzicz F, Jamieson T, et al. Artificial intelligence and the implementation challenge. *Journal of medical Internet research* 2019;21(7):e13659.
17. Whittlestone J, Nyrop R, Alexandrova A, et al. Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. London: Nuffield Foundation, 2019.
18. The Montréal Declaration for a Responsible Development of Artificial Intelligence., 2017.
19. Renda A. Artificial intelligence: Ethics, governance and policy challenges. *CEPS Task Force Report* 2019
20. Leslie D. Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. Available at SSRN 3403301: UK Office for Artificial Intelligence, 2019.
21. Nuffield Council for Bioethics. Artificial intelligence (AI) in healthcare and research. Bioethics Briefing Note, 1-8., 2018.
22. Whittlestone J, Nyrop R, Alexandrova A, et al. Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. *London: Nuffield Foundation* 2019
23. Machirori M. The Knotted Pipeline: Data-driven systems and inequities in health and social care: Ada Lovelace Institute, 2022.
24. Silverman D. Interpreting Qualitative Data 2011.
25. Tavory I, Timmermans S. Two cases of ethnography: Grounded theory and the extended case method. *Ethnography* 2009;10(3):243-63.
26. Glaser BG, Strauss AL. The discovery of grounded theory: Strategies for qualitative research: Routledge 2017.
27. ZOOM cloud meetings [program], 2022.
28. Sonix [program]. San Francisco, USA: Sonix Inc., 2022.

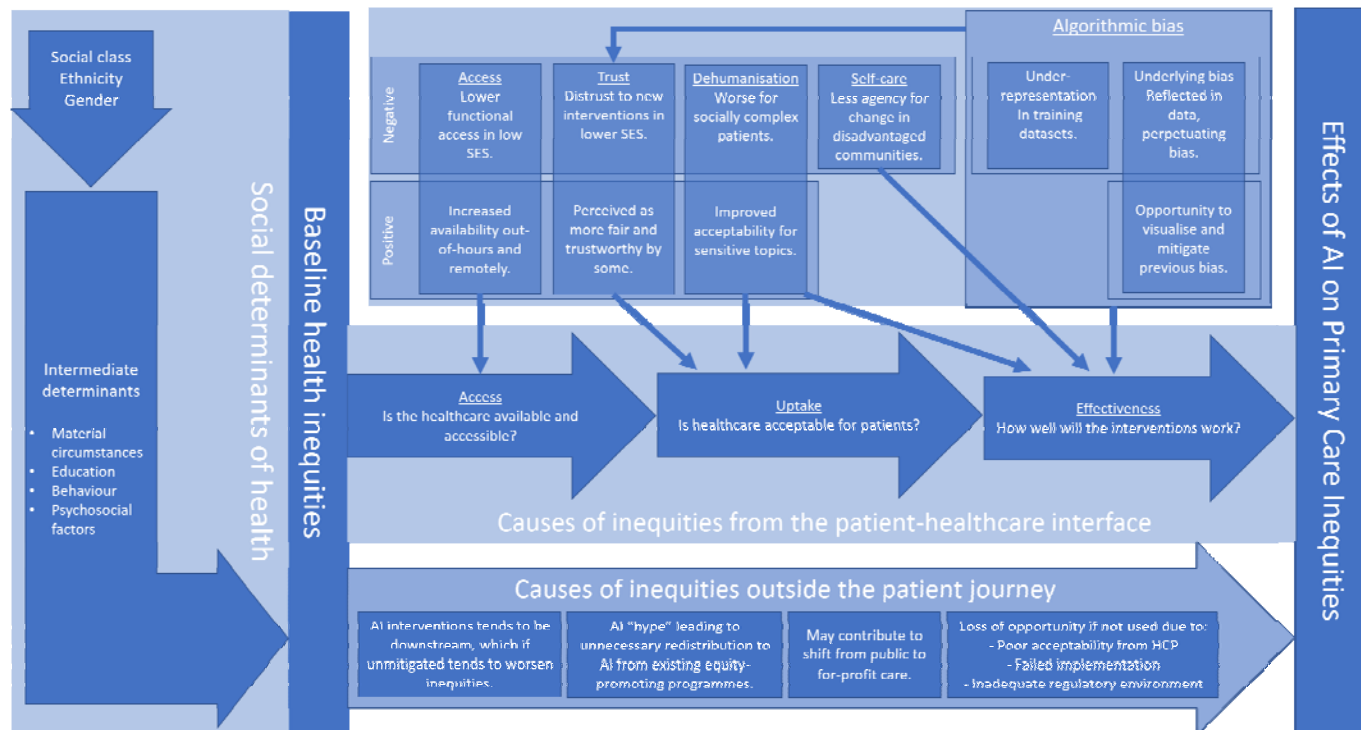
29. NVivo [program]. 12.0 version, 2020.
30. Ashcroft R, Menear M, Silveira J, et al. Inequities in the delivery of mental health care: a grounded theory study of the policy context of primary care. *International Journal for Equity in Health* 2021;20(1):1-11.
31. Knaak S, Patten S. A grounded theory model for reducing stigma in health professionals in Canada. *Acta Psychiatrica Scandinavica* 2016;134:53-62.
32. McCreaddie M, Lyons I, Watt D, et al. Routines and rituals: a grounded theory of the pain management of drug users in acute care settings. *Journal of clinical nursing* 2010;19(19-20):2730-40.
33. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and agenda for design justice. *Health Expect* 2021;24(4):1072-124. doi: 10.1111/hex.13299 [published Online First: 2021/06/13]
34. Lalani M, Baines R, Bryce M, et al. Patient and public involvement in medical performance processes: a systematic review. *Health Expectations* 2019;22(2):149-61.
35. Robert G, Cornwell J, Brearley S, et al. What matters to patients? Developing the evidence base for measuring and improving patient experience. *NHS Inst Innov Improv* 2011
36. Cadario R, Longoni C, Morewedge CK. Understanding, explaining, and utilizing medical artificial intelligence. *Nature human behaviour* 2021;5(12):1636-42.
37. Office for National Statistics. Population estimates by ethnic group and religion, England and Wales: 2019: ONS, 2019.
38. Greenhalgh T, Wherton J, Sugarhood P, et al. What matters to older people with assisted living needs? A phenomenological analysis of the use and non-use of telehealth and telecare. *Soc Sci Med* 2013;93:86-94. doi: 10.1016/j.socscimed.2013.05.036 [published Online First: 2013/08/03]
39. Frith L, Hepworth L, Lowers V, et al. Role of public involvement in the Royal College of Physicians' Future Hospitals healthcare improvement programme: an evaluation. *BMJ Open* 2019;9(9):e027680. doi: 10.1136/bmjopen-2018-027680 [published Online First: 2019/09/14]
40. Zidaru T, Morrow EM, Stockley R. Ensuring patient and public involvement in the transition to AI-assisted mental health care: A systematic scoping review and

- agenda for design justice. *Health Expectations* 2021;24(4):1072-124. doi:
<https://doi.org/10.1111/hex.13299>
41. Greenhalgh T, Wherton J, Papoutsi C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of medical Internet research* 2017;19(11):e367.
 42. Donia J, Shaw JA. Co-design and ethical artificial intelligence for health: An agenda for critical research and practice. *Big Data & Society* 2021;8(2):20539517211065248.
 43. Sugarman WB, Leila. Digital Health Hubs: An evaluation for the NHS Widening Digital Participation Programme: Good Things Foundation, 2020.
 44. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. *Sociology* 2009;43(3):535-54.
 45. Han YY, Carcillo JA, Venkataraman ST, et al. Unexpected increased mortality after implementation of a commercially sold computerized physician order entry system. *Pediatrics* 2005;116(6):1506-12.
 46. Ada Lovelace Institute. NMIP algorithmic impact assessment (AIA) template 2022 [Available from: https://docs.google.com/document/d/12HXv7Kb4dZLnA0BkL7DiccBxoq-Slg2meBsUBq_QQQI/edit#heading=h.y57c33w3aj9c accessed 2022-06-15.
 47. Public Health England. Health Equity Assessment Tool (HEAT). London: PHE Publications, 2020.
 48. NHS England and Local Government Association. Thriving places: Guidance on the development of placebased partnerships as part of statutory integrated care systems., 2021.
 49. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. <http://www.nice.org.uk/about/nice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp> 2008
 50. NHS Digital. GPIT Futures 2022 [Available from: <https://digital.nhs.uk/services/gp-it-futures-systems>.
 51. NHS Digital. GP IT Futures buying catalogue: procurement process 2022 [Available from: <https://digital.nhs.uk/services/gp-it-futures-systems/beta---buying-catalogue---procurement-process>.

52. NIHR ARC NWC. Health Inequalities Assessment Toolkit (HIAT), 2017.
53. World Health Organization. Innov8 approach for reviewing national health programmes to leave no one behind: technical handbook: World Health Organization 2016.

Appendix 1

Conceptual framework model for how AI could affect inequities in primary care, forming part of the theoretical foundation for this study. From d'Elia et al (2022).



Appendix 2

Schematic implementation process for AI in English NHS primary care, based on cited documents and fact-checked with interviewees:

1. Development:
 - a. A producer e.g. a private company, NHS organisation, academia, or any combination thereof (but most likely a private company) develops a product they believe is suitable for primary care.
 - b. The producer may or may not consult guidance on how to create an equity-improving AI product, for example through the abovementioned HEAT, through an AIA such as the one by The Ada Lovelace Institute⁴⁶, NICE ESF⁴⁹ or the guidance issued by The UK Office for Artificial Intelligence and The Alan Turing Institute²⁰.
2. Approval:
 - a. Producer seeks MHRA approval. MHRA produces standards according to UK MDR 2002 legislation. These standards are applied by accredited third-party organisations, which provides MHRA approval for the product. If it is a class 1 product (only advising) the producer can self-assess and provide documentation, otherwise the MHRA must be involved in the clinical assessment of efficacy.
 - b. If the system is intended to be used for traditionally primary-care-supplied services but not be used in actual NHS primary care (e.g. private healthcare or online symptom checkers), no further assessment is necessary.
 - c. If produced together with a local partner such as a university or a PCN, an AHSN or ICB may be involved in this process.
 - d. The producer may apply for the product to be placed in the GPIT Futures catalogue of endorsed software. This is not an absolute requirement for being procured into the NHS primary care system, but advised⁵⁰.
3. Procuring:
 - a. An ICB or a PCN (or rarely, a single primary care practice) sees a use case for the product and investigates procuring it. An AHSN is likely involved with the ICB in the case of ICS-wide employment. The ICS and the AHSN may have a PPI component advising the ICB on if and how to implement the product. How exactly the PPI is shaped varies between organisations, but it is expected (if not statutory) for both ICBs and AHSNs.
 - b. If the buyer is a PCN they would be expected, but not required, to ask the ICS and associated AHSN for support.
 - c. If under £40,000 and under twelve months contract time the buyer can self-procure, otherwise it has to be a competitive process between the various products meeting the specification in the GPIT Futures catalogue⁵¹.
4. Operationalisation:
 - a. NHS Digital publishes guidance on operationalisation of new products. Their guidance is strictly operational in character and does not cover PPI/needs assessment and feedback from patients or end-users⁵¹. However,

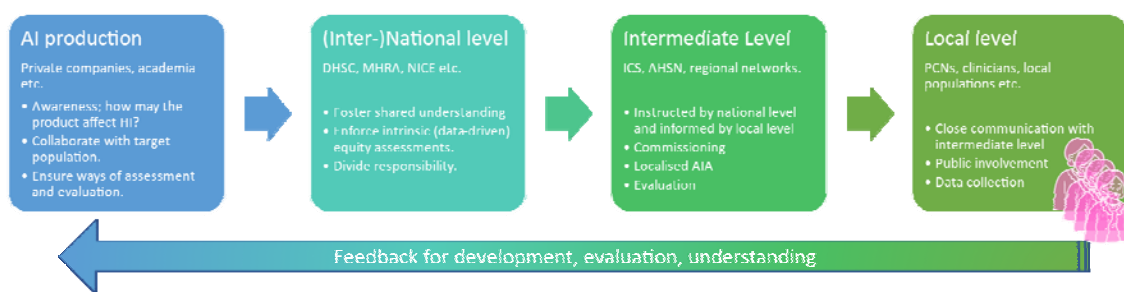
it is assumed that those aspects have been covered to an extent by the ICB or associated AHSN.

5. Evaluation:

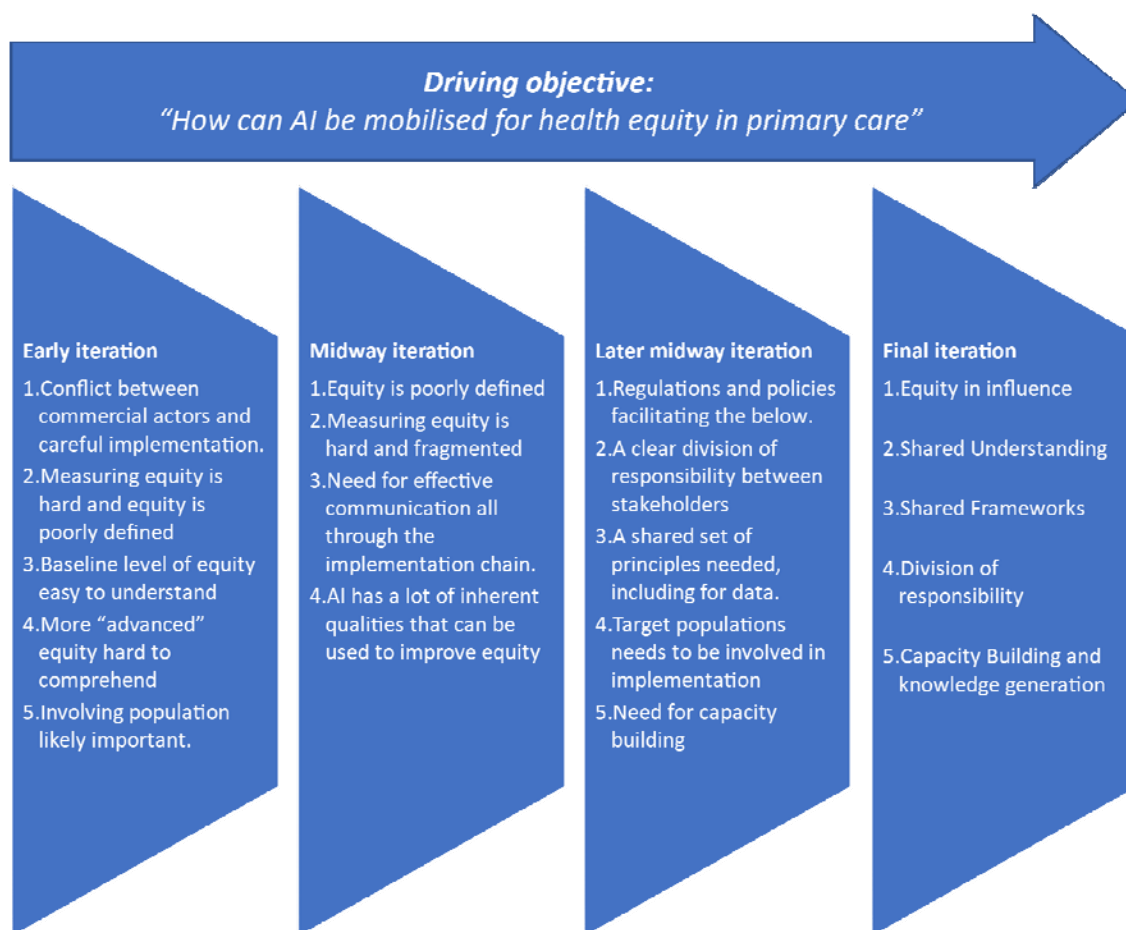
- a. There are no detailed guidelines for evaluation, and no equity-specific evaluation is regularly done, but the buyer would in practice evaluate the products against the objectives it was brought in under. NHS Digital recommends annual evaluations⁵⁰. The CQC has a responsibility to assess whether products are *used* in a safe and effective manner from a clinical perspective but does not evaluate the effectiveness and other properties of the products themselves (e.g. algorithmic biases).

Appendix 3

Based on the mapping in appendix 2 we combined fundamentals of implementation science (e.g. the need for a shared understanding of goals, benefits and measurables of a given intervention ⁴⁴) and general approaches to inequity improvement in healthcare (such as HEAT ⁴⁷, inequityAT ⁵² or Innov8 ⁵³, among others). Our resulting a conceptual model for equitable implementation was developed to show how equitable implementation is possible. Here we illustrate the implementation chain that we recommend is followed from production to utilisation:



Appendix 4



Sample of iterations of themes derived from coding interview data. Note that multiple minor revisions were made between these versions.

Table 1

Table 1: Overview of individual interviews and group interviews with people living with diabetes (PwDs). Note that age and ethnicity were only recorded for the latter. Numbers refer to anonymization code, broadly corresponding to the order of recruitment.

Professional stakeholders	% female	Professional stakeholders	% female
Total subjects (excl. focus groups) n=27	36 %	Commercial developers n=4	25 %
		19	
General Practitioners n=7	57 %	20*	
1		22	
2		23	
3			
5		Commercial developers: Focus group n=1	(33 %)
7		21 (n=6)	
11		Local commissioners (ICS) and Advanced Health Service Networks (AHSN) n=4	50 %
16		6 (ICS)	
		8 (ICS)	
Regulatory/advisory stakeholders n=6	33 %	13 (AHSN)	
4 (Dept. Of Health and Social Care)		15 (AHSN)	
10 (Medicines and Healthcare products Regulatory Agency)			
12 (Information Commissioners Office)		Researchers n=6	17 %
17 (The Ada Lovelace Institute)		9	
18 (Care Quality Commission)		14	
27 (National Institute for Health and Care Excellence)		24	
		25	
		26	
		28	
PwDs: Focus groups n=4	% female	% >65 years	% minority ethnicity
PwD 1 (n=7)	57 %	100 %	29 %
PwD 2 (n=6)	33 %	83 %	50 %
PwD 3 (n=6)	67 %	100 %	0 %
PwD 4 (n=10)	60 %	70 %	10 %
Total PwDs: n=29	55 %	86 %	21 %
Average duration of interviews: 39 minutes (range 23-63). Focus groups: 43 minutes (33-51)			

Figure 1

Title: The iterative process for data analysis and recruitment.

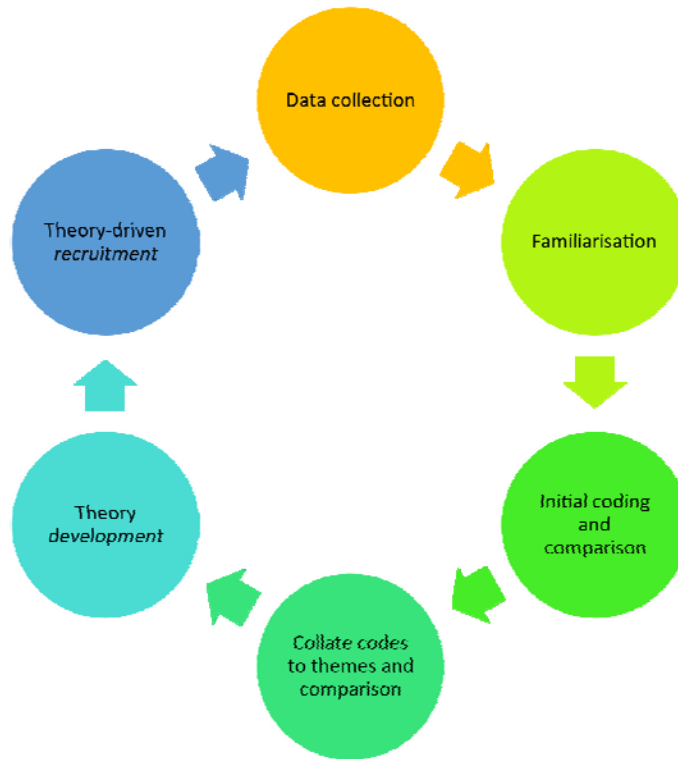


Figure 2

Title: The theory model.

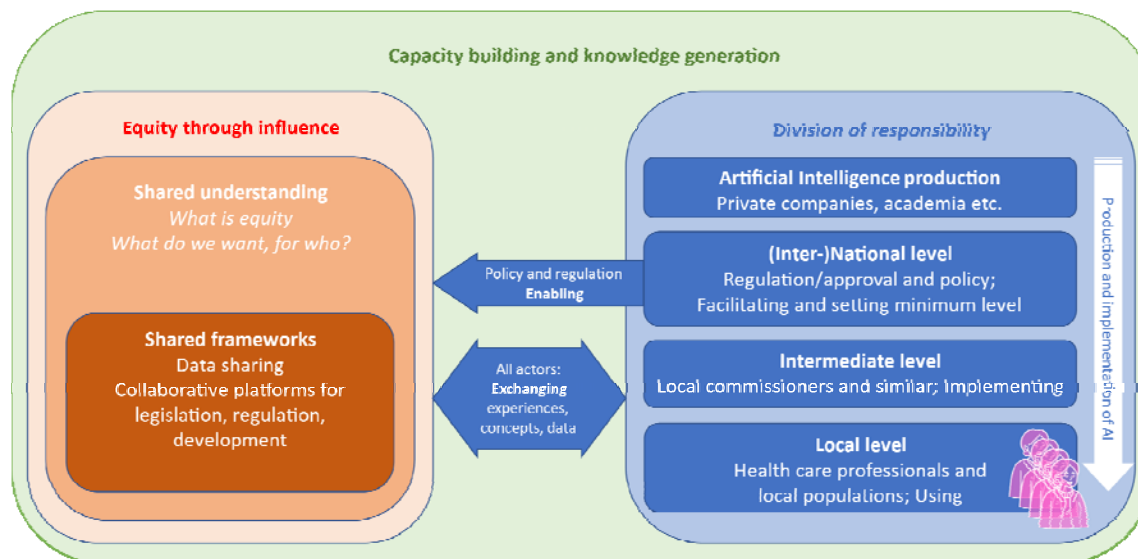


Figure captions

Figure 1: The iterative process for data analysis and recruitment. 1) Data collection, 2) Familiarization of transcripts and field notes, 3) Initial coding and comparison, 4) Collate codes into themes. 5) Build provisional advance theory from themes. 6) Recruit new participants to explore the themes and theory further. Reiterate as new data are collected until saturation.

Figure 2: The theory: the concluding model for equitable implementation of AI in primary care.