Health Equity in Clinical Research Informatics

Sigurd Maurud, Silje H. Henni, Anne Moen
University of Oslo, Oslo, Norway

1 Introduction

Clinical research informatics (CRI) is a sub-discipline within biomedical and health informatics that focuses on the analysis, interpretation, and presentation of clinical knowledge generated through informatics [1]. This definition by Embi and Payne [1], dates back over a decade ago and has previously been mentioned in this journal by Solomonides [2]. Acknowledging and including topics that have flourished since Embi and Payne’s definition [2], the most notable addition is Artificial Intelligence (AI), referring in a broad sense to the ability of technology to resemble functions and processes of human beings. Machine Learning (ML) [3], and Natural Language Processing (NLP) [4, 5], are notable subgroups of AI technologies using relevant clinical data sets to advance the representation and understanding of a problem. Increased availability of clinical data in digital form and expanding computational capacity enable more complex and sophisticated processing of clinical data in CRI [6]. This appears as prevalent use of AI techniques in CRI, seeking to advance clinical practice through decision-support and prediction capabilities to support health practitioners in different specialties [7-14]. However, extensive use of AI algorithms has also revealed potential risks with implications for patients and their prospects of best possible treatment [15]. An example is the “black box” challenge, which delineates the necessity of making complex AI operations transparent and comprehensible to end-users [16]. Vast databases are being queried by algorithms, as researchers and clinicians are seeking patterns that can guide decisions in clinical care and result in opaque explanations for how the algorithms reach their guidance [17]. The challenge here may not be a result of secrecy or inadequate knowledge, but rather ML outputs without regard for human comprehension and careful consideration of clinical relevance [18]. In other words, “black box” approaches to decision-making in patient care may incur significant risk, as neither practitioner nor patient can fully comprehend the steps leading to recommendations [3, 19]. In addition to well-known problems being addressed in the development of AI services in health [20], issues like access to and ownership of clinical data, and possible exacerbation of health inequity [21, 22] are important ethical issues of concern. The focus on these topics seems further accelerated by the unprecedented deployment of digital solutions within healthcare during the COVID-19 pandemic. Consequently, several issues were illuminated, including dependence on digital health, how to enable digital solutions to provide better healthcare, as well as existing inequalities and structural discrimination [23].

Health is associated with a non-medical social gradient, where those on the lower socioeconomic end often have the least chance for good health. The circumstances in which an individual is born, ages, lives, and works, usually referred to as “social determinants of health” (SDoH), can be exacerbated by discrimination, prejudice, and stereotyping [24]. Health equity, as defined by the World Health Organization, is the absence of unfair, avoidable and remediable differences in health status among groups of people, and requires actions that even out differences in health outcomes between populations with different socioeconomic foundations [24]. While holding great promise for the use of AI in health care, CRI can pose risk of reflecting and reproducing analytical and algorithmic biases that potentially increase health inequities that come with SDoH [21, 25]. Algorithmic bias that discriminates based on characteristics integral to the
person, such as race and ethnicity, has received particular attention in this context [26-30]. Considering ‘race’ and ‘ethnicity’, ‘ethnicity’ comes with similar interpretation in the literature, but it is necessary to address the conflicting use of the term ‘race’ in European and American contexts. The US Census Bureau and the Office of Management and Budget (OMB) refer to race as a socially constructed way of separating humans into different sociocultural and ancestral groups [31], while in Europe, the term ‘race’ is avoided due to its association with the wrong notion of biological differences among human beings, preceding the linked historical and ideological associations. Racism is, on the other hand, an acknowledged term in Europe, referring to discrimination based on the notion of biological differences [32]. Recognizing this distinction, this article uses the term ‘race’ when referring to the cited sources that apply this in line with the definition of the US Census Bureau and OMB.

A key concern in Real World Data (RWD) based studies is representativeness. Using such data sets for training algorithms poses a risk for algorithmic biases in AI [33], originating from, e.g., lack of inclusion of underrepresented population groups in samples [34, 35], and subjective assessments [36, 37] within the data material. An example of this problem is the socially inconsistent and intermixing use of the terms ‘ethnicity’ and ‘race’, possibly affecting the creation and collection of data [30]. Another key concern is to delineate which circumstances one is to discern between different populations as it may be of relevance to some conditions and completely irrelevant to others [38].

CRI hold much promise for improving clinical practice [1], but needs to incorporate assessment of its impact on health equity to provide healthcare to all patients, regardless of SDoH [16, 21, 39]. Illustrating this concern is the establishment of the High-Level Expert Group on Artificial Intelligence (AI HLEG) by the European Commission [15], to accommodate for the implementation of the Commission’s vision for ethical AI [40]. As an output, seven requirements for Trustworthy AI have been published from this group: 1) Human agency and oversight, 2) Technical robustness and safety, 3) Privacy and data governance, 4) Transparency, 5) Diversity, non-discrimination and fairness, 6) Societal and environmental wellbeing, and 7) Accountability [15]. Taking this into account, CRI should go beyond monitoring, controlling, and guarding against unintentional outcomes that may exacerbate structural health inequality, to actively address and hence, improve health equity [21, 35, 41]. Experience gained during the COVID-19 pandemic has highlighted the need for a more systematic approach to ensure that digital health and CRI promotes health equity and the goal towards universal health coverage [23]. With this aspiration and inspired by the topic of the 2022 IMIA Yearbook: “Inclusive Digital Health: Addressing Equity, Literacy, and Bias for Resilient Health Systems” [42], the aim of this scoping review was to examine in what ways research in CRI, published in the year of 2021, has included health equity to promote patient health and care.

2 Method

This scoping review applied methods as outlined in the Joanna Briggs Institute (JBI) Manual [43]. The process of the review proceeded as follows: 1) identify research question, search terms and keywords, 2) search for literature, 3) screening and selecting relevant literature, 4) extract data from selected literature and, 5) summarize and present the results. A protocol defining the research question, aim, screening process, search terms and criteria for inclusion and exclusion was developed in advance of the literature search. The approach is illustrated in a PRISMA flow diagram (see Figure 1) [44].

2.1 Search Strategy

A medical librarian guided our search in September 2022, using the following databases: Medline, Embase, ACM library and Epistemonikos. In line with the JBI Manual [43], a PCC framework was applied for the literature search:

2.2 Screening and Selection of Literature

Ahead of the screening process, we screened the titles and abstracts of 25 randomly selected sources from the search results to reach a general agreement on Inclusion and Exclusion criteria before the selection of sources (see Table 1). All sources were then screened by the first and second author using the predetermined criteria for inclusion and exclusion. The sources were screened in two subsequent rounds supported by Covidence, a web-based collaboration software platform used for screening and data extraction in literature reviews [45]. The first round of screening extracted literature based on titles and abstract, while the second round of screening extracted literature through full-text reading. The first round of screening resulted in conflicts on 86 sources (18% of the total sources screened), all of which were resolved through a plenary review. The second screening round resulted in one conflict among the 58 sources that underwent full-text review. The conflict was resolved through a plenary review of the source. Specific quality assessment of the literature was not carried out as this is generally not a priority in scoping reviews [43].

2.3 Data Extraction

A spreadsheet with the data material was created to extract information on the study reference, population characteristics and key findings that relate to the aim of the scoping review. The first and second authors read the full text sources with the purpose to identify and extract aspects of clinical research informatics, aspects of health equity, and aspects of patient implications.
3 Results

Of the 58 sources that underwent full-text review, eight studies were included in this study. The reasons for exclusion are listed in Figure 1. Although five sources focused on health equity in CRI, they did not focus on what ways CRI can drive health equity and was therefore excluded under the reason “General data suitability and ethical considerations for AI research”. Among the eight included studies, three were reviews and therefore controlled for duplicates. Patra et al. [46] and Pham et al. [47] are both citing Hazlehurst et al. [48]. Patra et al. [46] also share an article with Craig et al. [49], both citing Navathe et al. [50]. However, as the included sources have different focus and have contributed different findings, we did not consider the reviews as overlapping. No overlapping articles were identified between Pham et al. [47] and Craig et al. [49]. With the exception of one Canadian study [47], all included studies were conducted in the United States. The focus on CRI in these papers was consistently on AI technology with two themes identified on how CRI with patient implications can drive health equity:

1) Exposing health inequity in CRI and addressing the need for adequate measures;
2) Promoting health equity through CRI.

The included publications are presented in Table 2 with descriptions of how health equity in CRI with patient implications is addressed in each paper.

4 Discussion

Health equity, CRI and AI are topics in a global setting. It is therefore interesting that all the included papers in this review are of North American origin. Except for one Canadian study, all included studies are conducted in the US. It may be that this simply reflects a greater focus on health equity in CRI and AI in the US compared to other countries, responding to a policy that focuses on promoting equity and justice for all [57, 58]. Furthermore, low- and middle income countries still appear to face challenges considering the implementation.
<table>
<thead>
<tr>
<th>Reference</th>
<th>Health Equity</th>
<th>Clinical Research Informatics</th>
<th>Patient implications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coley et al., [51]</td>
<td>The study evaluated socio/ethnic differences in the performance of two prediction models for death by suicide.</td>
<td>The object under assessment was a risk algorithm, trained on health record data, with the aim of predicting suicide within 90 days after an outpatient visit in mental healthcare.</td>
<td>Health record data may be inadequate for the prediction of suicide in some racial/ethnic groups, due to insufficient data on death by suicide for racial/ethnic groups with low prevalence, but also systemic barriers, practitioner bias, institutionalized discrimination, outcome misclassification, and cultural differences. Currently accessible methods do not address subgroups of the entire sample. Clinical prediction models need to be assessed for performance within the population group of relevance to bridge the gap to underserved populations.</td>
</tr>
<tr>
<td>Craig et al., [49]</td>
<td>Though a literature review, this study examined how SDoH data and health technology are utilized to enhance the management of population health.</td>
<td>AI holds the possibility to increase the acquisition and utilization of SDoH data from electronic health records (EHR).</td>
<td>The study concluded that technologies with advanced analytical methods holds promise for considerate, inclusive, and effective personalized care and may provide positive contributions to decision-making processes.</td>
</tr>
<tr>
<td>Hammarlund [52]</td>
<td>The study examines whether differential health factors can explain racial treatment disparity in cardiac surgery.</td>
<td>Conventionally, the applicability of health risk disparity in treatment disparity is tested with adjusted observed rates for patient-level health differences to optimize patient comparability. The study suggests an alternative, using a ML-based method that enhances the capture of clinical practice to adjust for the critical predictors in acute myocardial infarction surgery.</td>
<td>The suggested ML-based method decreased the standard adjusted acute myocardial infarction surgery disparity by 45-59%. After adjusting for predictive controls, a smaller, but still present surgery disparity remains, indicating a differential racial treatment explained by other factors than health risk differences. Physician experience with treating black patients appeared to reduce stereotypical statistical discrimination.</td>
</tr>
<tr>
<td>Lu et al., [53]</td>
<td>With the aim of improving equity in healthcare access throughout the COVID-19 pandemic, the study developed and tested a ML-based model that matched predicted patient language with contact tracer language in communities highly represented with language barriers.</td>
<td>The ML-based model used random forests, integrating laboratory reports data with census data to predict the language of any incoming patient. To ensure interpretability for the stakeholders, the study based the model on a small feature set, optimizing transparency and comprehensibility.</td>
<td>The study demonstrates how a small ML-based measure can enhance effectiveness of infectious disease response and at the same time support health equity by matching cases to contact tracers based on language.</td>
</tr>
<tr>
<td>Patra et al., [46]</td>
<td>The review examined different approaches for indexing and providing SDoH data through natural language processing (NLP).</td>
<td>Matching of keyword and classification were the most prevalent approaches to identify SDoH, either through supervised or unsupervised algorithms. Due to insufficient data quantity for certain SDoHs, the review reports lower performance of neural networks compared to traditional ML-algorithms.</td>
<td>The study promotes an NLP-based SDoH identification to tackle the increasing acknowledgment of nonclinical factors’ influence on health. The study argues that NLP-based methods to index and develop outcome analysis in SDoH may optimize clinical decision-making without greater interference in current workflows.</td>
</tr>
<tr>
<td>Pham et al., [47]</td>
<td>The potential for AI technology to be algorithmically biased in disfavor for underserved populations led this case study to review and propose ethnoracial considerations in AI-based diabetes management tools.</td>
<td>The authors express particular concern about distribution shift in AI models, discrepancy between the training population and the end-users leading to inaccurate predictions.</td>
<td>The authors identified a small degree of ethnoracial data in the reviewed articles (10 out of 141 articles on AI-based diabetes tools). The study calls for measures that facilitate inclusive algorithms that provide diversity consideration of healthcare innovations.</td>
</tr>
<tr>
<td>Pierson et al., [54]</td>
<td>Through a ML-based method, the study demonstrates how standard radiographic measures of osteoarthritis severity ignore undiagnosed features that inordinately affect underserved populations with knee pain.</td>
<td>As an alternative to the standard radiographic measure Kellgren-Lawrence Grade (KLG), the study trained a convolutional neural network on a dataset of radiographs to predict reported pain score for each knee. Predictions were generated in an independent validation set separated from the training set.</td>
<td>The algorithmic pain prediction measure accounts to a larger degree for disparities in osteoarthritis compared to KLG and is proposed by the authors as a tool to enhance access to treatment for underserved patients with osteoarthritis.</td>
</tr>
<tr>
<td>Thompson et al., [55]</td>
<td>The study assessed the predictions of an existing NLP opioid misuse classifier, using principles of fairness, accountability, transparency, and ethics to tackle the possible impact of historical and structural bias on disadvantaged subpopulations.</td>
<td>The study adapted a python toolkit for reviewing fairness and bias in ML-models and the Local interpretable Model-Agnostic Explanations [56] to assess face validity. The post-hoc results demonstrated disparities in patient characteristics across subgroups.</td>
<td>The authors argue that collaboration with data scientists and ethic experts in advance of model training, could have addressed disparities and biases in the model, although inequalities present in the model may only be addressed in post-hoc recalibrations. The study demonstrates the result of disproportionately overlooking potential biases in the validation of convolutional neural networks.</td>
</tr>
</tbody>
</table>
of AI in health [59]. The US, compared to other high-income countries, are ranking well below in health care system performance. Among the domains that pull the performance down are less access and equity, despite the significant amount of the US gross domestic product that are spent on health care [60].

Although it appears the COVID-19 pandemic has highlighted established disparities in health and digital health [23], interestingly only one study identified in the survey as published in 2021 acknowledged the COVID-19 pandemic as a catalyst [53]. However, the recognition of social inequities and its influence on health or digital health in all included publications may rather reflect the COVID-19 pandemic’s role as a catalyst in illuminating and driving the focus on health equity [23].

4.1 Exposing Health Inequity in CRI
In line with concerns articulated in previously cited literature [21, 25-30], three of the included sources discussed risk for health inequity in AI-based health technology [47, 51, 55]. Coley et al. [51] assessed differences in the performance of two prediction models. They found that these addressed subpopulations of ethnic or racial minorities in an inadequate manner, identifying a smaller proportion of anticipated suicides in patients who report race/ethnicity to be Black and Alaskan Native/American Indian, and in patients who do not report race/ethnicity. The secondary analysis conducted by Pham et al. [47] of 141 articles from a literature review, looked into ethnoracial considerations in AI diabetes tools, to propose a strategy for equity for such technology. As the creators of an NLP opioid misuse classifier, Thompson et al. [55] evaluated the impact of bias against historically and structurally disadvantaged groups. All three CRI studies acknowledged the challenges for health equity within AI, mainly expressed through algorithmic bias [47, 51, 55]. This is echoing the fact that RWD used to train algorithms risk reproducing bias in technological solutions [33], possibly through lower accuracy for underrepresented samples of underserved groups [34, 35], or through subjective assessments [36], reinforcing possible judgemental biases from healthcare providers [37]. Thompson et al. [55] acknowledge their previous lack of consideration for disadvantaged populations in the creation of their instrument, and thus mirror the concern of Coley et al. [51] for insufficient attention to the clinical usefulness or utility of AI technology to disadvantaged subpopulations. Besides, Pham et al. [47] identified only 10 out of 141 papers on AI diabetes tools that inconsistently addressed race or ethnicity, or both (race/ethnicity), pointing to a lack of reliable data and a lack of focus for ensuring adequate training algorithms for ethnic or racial minority populations. Even if assessed for algorithmic bias, the “black box” nature of AI will still challenge transparency, potentially including unintended bias, or withhold information underlying the performance of a model [17, 18, 55]. This is crucially important [21, 22], as it emphasizes the responsibility of CRI to acknowledge and act upon this in digital health technology [46, 55], and incorporate principles for ethical AI, as outlined by the AI HLEG [15].

4.2 Promoting Health Equity in CRI
The remaining five papers examined ways in which CRI may enable and promote health equity. Craig et al. [49] and Patra et al. [46] did so in an indirect manner, through literature reviews examining and promoting the utility of AI to actively include and use SDoh data from electronic records. However, it appears to be beyond the scope of both reviews [46, 49], to discuss value in subjective data, as well as potential bias introduced by the source of data. As clinical text include subjective data in EHR [36], this illustrates the issue of possibly overlooked subjective bias in algorithmic performance [30, 36, 37]. Indeed, algorithmic bias appears in general as a difficult barrier for health equity to overcome [26-30, 33-37, 51, 55]. The current conditions where CRI demonstrates promotion of health equity appears admittedly to be those where CRI is used specifically to address inequity in health; not only in the evaluation of AI-based healthcare instruments ability to promote health equity [51, 55], but through AI-based methods demonstrating and addressing disparities in the delivery of healthcare services [52-54].

Building from observations that underserved populations experience a greater amount of pain in osteoarthritis, Pierson et al. [54] used a deep learning approach on radiographs to predict the pain level of the individual patient, finding that the approach significantly reduced unexplained racial pain disparities compared to traditional methods. Through a ML-based method, Hammarlund [52] demonstrated disparities between black and white patients in acute myocardial infarction treatment, beyond that explained by health risk differences. As a response to how language discordances limited the contact tracing of a non-English speaking population in California, already disproportionately affected by COVID-19, Lu et al. [53] used an ML-based approach to predict the language of an incoming patient and match this to the language of the contact tracer. In contrast to the other included sources in this scoping review, these sources address health equity by directly responding to existing health inequities. Pierson et al. [54] and Hammarlund [52] both have in common the use of AI in exposing health inequity in clinical practices and providing alternative solutions, while Lu et al. [53] uses AI to promote health equity in a setting known to be characterized by disparities in health and access to health. All the included sources of this scoping review address health equity in CRI with patient implications, either by exposing health inequality in AI-based solutions or by examining possibilities for AI to extract data of importance to address health equity [46, 47, 49, 51, 55]. However, Pierson et al. [53], Lu et al. [53], and Hammarlund [52] all stand out in their application of AI to drive health equity. The accomplishments of these three studies appear to stem from how they approach the issue of health equity. Instead of illuminating health inequity present in CRI-driven solutions, such as algorithmic bias within AI-based prediction models [51, 55], they use AI to promote and improve health equity in the deliverance of existing treatments and health care services [52-54].
4.3 The Way Forward

To further assess the result of our findings, we performed a similar search for the year of 2022 to discern if more recent literature would add to the significance for this study. We identified at least 21 papers [61-81] that met our inclusion criteria, including results from the 2022 IMIA Yearbook [64, 68, 78]. Reading through these articles, we did not identify additional thematic areas than those we have included for the year of 2021. However, it appears to be a change in terms of attention to the topic. The focus on health equity in CRI seems to increase considering the 21 studies we found published in 2022 compared to the 8 from 2021. Following this again, three studies were also identified, just for the first month of 2023 [82-84]. The interest for the topic are expanding from North America based study reports to other parts of the world, including Europe [63, 65, 67, 77] and Asia [84]. The scope of health equity in CRI also appears to have expanded and evolved. Primarily centred on challenges considering race and/or ethnicity in 2021 [47, 51-55], health equity in CRI has extended to diagnosis bias in rural populations [75], age [64], and gender or sex-specific bias [63, 64, 67, 70, 75, 77].

5 Conclusion

Several of the studies on Clinical Research Informatics presented here highlight algorithmic bias as a factor in the promotion of health equity in digital solutions [47, 51, 55]. It appears to be a considerable challenge for CRI to provide AI-based solutions without algorithmic bias that prove counterproductive to the intention and goal of the solutions. Carefully selecting and appropriately balancing different characteristics may reduce algorithmic bias and adjust outcomes in some cases, but bias can also remain hidden which make correction nearly impossible [38]. Based on the findings in this scoping review, our impression is that the field of CRI, here exemplified by AI as the focus of the recent publications found, is more aware of the challenges at hand, which is an important starting point to find solutions that remedy this challenge. This way CRI will increase capability to promote and improve health equity. This review illustrates that when the right form of digital technology is correctly adapted to the population in question at the right time, AI-based CRI-solutions hold a promise to drive equity in health. Recent publications, in 2022 and beyond, illustrate advancements and endeavour to improve AI algorithms that leverage and combine efforts to reduce and eliminate algorithmic bias.

Further progress and full incorporation into CRI require thorough assessment and improvement for equitable and ethical distribution of health care services that respect patient autonomy and dignity. Going forward, CRI holds opportunities for novel patient- focused digital tools that stimulate engagement and promote health equity. This requires tools that do not exacerbate structural inequalities, incorporate ethical consideration to avoid harm, and mitigate risks related to sub-populations already exposed to disparities in society and health.

Acknowledgement

We would like to thank Toril M. Hestnes, Senior Librarian, University of Oslo, Library of Medicine and Science, for her contribution to the literature search. The work is partly funded by University of Oslo (AM), Gravitate-Health, EU H2020 agreement 945334 (SM), and CORAL, the Research Council of Norway project 301517 (SHH).

References


65. D’Elia A, Gabbay M, Rodgers S, Kierans C, Jones E-mail: sigurd.maurud@medisin.uio.no


