The Use of Artificial Intelligence in Clinical Diagnostics – Challenges To Consider For Implementation

PÁDRAIG CRONIN

Introduction

Whilst many technological advancements have revolutionised healthcare throughout the 21st century, one of the most significant is Artificial Intelligence (AI). AI is generally regarded as the capability to imitate intelligent human behaviour using machines, and is based on computer science, statistics, algorithms, machine learning, information retrieval, and data science[1]. AI has permeated into many domains of healthcare including Clinical Diagnostics. While AI chatbots (such as those used in Babylon and Ada) are being used by patients to identify symptoms and recommend further actions in community and primary care settings, more recent advances in the technology with larger datasets have provided users access to a more extensive array of clinical conditions[2].

However, as these tools are constantly being developed with an ever-increasing dataset of clinical cases, certain challenges threaten the implementation of an accurate and effective model. In this article, the issue of Data Bias, and Data Handling will be examined within the context of Clinical Diagnostics, and how these factors threaten the development of such AI Healthcare tools.

Data Bias - An Unrepresentative Algorithm

Data Biases are a fundamental threat to the development of AI Tools which arise from errors within the data collection process. These biases can produce harmful results for people, including social discrimination and a significant loss of trust from society, if they are not identified and mitigated[3]. This is particularly true for medical diagnostics where a wide range of signs, symptoms, laboratory results, and clinical presentations are used across variable age ranges, genders, and ethnic backgrounds to determine an accurate diagnosis.

One of the most important steps in the development of any AI Tool is the cultivation and development of a suitable dataset with which to train the model upon. This dataset will form the basis of the algorithm that will ultimately assign a diagnosis. When developing such a dataset, many variables must be examined in order to create a fair and balanced outcome. Errors in the collection of clinical data include, but are not limited to:

- Unrepresentative Case Distribution – Creating a skewed probability output which favours certain population cohorts, and thus misdiagnoses minorities within the population.
- Vague or Irrelevant Signs/Symptoms – Drawing correlations between symptoms and diagnoses which may not be linked, and can result in model confusion and/or reduced accuracy.
- Incorrect Signs/Symptoms Input – Generation of confusion within the model based on a mismatch between symptoms and diagnoses.
- Reuse of Clinical Cases – Potential to bias the model toward a particular diagnosis, overrepresenting a particular condition.

In the event of any of the above circumstances, the final AI algorithm may be unrepresentative of the actual patient cohort being diagnosed, and is liable to produce inaccurate or false diagnoses. To combat this internal bias, each source of error must be carefully examined when reviewing the dataset prior to model development, and the process repeated with the addition of new data throughout the iterative training process.

https://doi.org/10.33178.SMJ.2024.1.12

Data Handling – Balancing Patient Privacy With Model Development

A major risk which poses a threat to the development of AI Clinical Diagnostic tools is the associated data handling during model development. A critical step in the development of any artificially intelligent model is the input of real data which can be used to drive accuracy and precision of the diagnostic outcome.

Many open-source AI tools collect the data which input by the user for two reasons: model development, and retesting of the model during iterative training. In order to carry this out within the context of a clinical diagnostic AI tool, patient data such as age range, ethnic background, signs/symptoms, and diagnosis would need to be accessible and stored for use. An example of this is the OpenAI tool, ChatGPT, which states in its guidelines that "Your conversations may be reviewed by our AI trainers to improve our systems"[4].

This presents with it the significant challenge of how to balance model development with patient privacy. Large scale data privacy laws which have been implemented within the past decade include the European Union's General Data Protection Regulation (GDPR) Article 17, which holds the "right to be forgotten" as a fundamental pillar[5]. This right entitles patients to have their data withdrawn from the databases which are being used for AI development. With this arises the conflict between developing an accurate model whilst accounting for patient privacy rights.

A possible solution to this would be the anonymisation and redaction of patient data which could then be used to drive model development. This process would fulfil the GDPR Article 17 requirements if a patient's identity cannot be recovered from the anonymised data[6]. This would require significant resources and procedures within the companies developing these AI tools. However such anonymisation would guarantee patient safety whilst preserving data that is critical for product development.

Conclusion

AI Toolkits are rapidly becoming an everyday feature of many disciplines, including healthcare. As demonstrated by its extensive use in other areas of industry, AI is an incredibly powerful tool that has the potential to transform healthcare in the future. However, it is not a perfect tool in its current form for a number of reasons as discussed above. Two areas of improvement include:

- Targeted and careful dataset design for algorithm development in order to mitigate diagnostic bias across various patient cohorts and disease profiles.
- Development of concrete and executable guidelines within the software training process which maintain patient privacy whilst also protecting model development.

Once both of these areas have been examined and addressed, AI not only has the potential to improve efficiency in the clinical environment, but also to improve patient diagnostics through the implementation of highly accurate clinical decisions driven by real world data.

References

1. AI Yao L, Zhang H, Zhang M, Chen X, Zhang J, Huang J, et al. Application of artificial intelligence in renal disease. Clinical eHealth. 2021 Jan 1;4:54–61.
2. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc J. 2021 Jul;8(2):e188–94.
3. European Parliament. Directorate General for Parliamentary Research Services. Auditing the quality of datasets used in algorithmic decision-making systems. [Internet]. LU: Publications Office; 2022 [cited 2023 Dec 20]. Available from: https://data.europa.eu/doi/10.2861/98930
4. What is ChatGPT? | OpenAI Help Center [Internet]. [cited 2023 Dec 20]. Available from: https://help.openai.com/en/articles/6783457-what-is-chatgpt
5. Union PO of the E. Publications Office of the EU. Publications Office of the European Union; 2016 [cited 2023 Dec 20]. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). Available from: https://op.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/language-en
6. Is the anonymization of personal data the same as data erasure? [Internet]. [cited 2023 Dec 20]. Available from: https://www.ironmountain.com/resources/blogs-and-articles/i/is-the-anonymization-of-personal-data-the-same-as-data-erasure