RAPID EVIDENCE SUMMARY

DELIVER
Research for better healthcare closer to home

How is artificial intelligence/machine learning being used in diabetes care?

Authors

Dr Olivia Metcalf, Senior Research Fellow, Centre for Digital Transformation of Health, University of Melbourne

Dr Kimberly Cukier, Endocrinologist, Barwon Health

Background

Currently, 1.9 million Australians are affected by diabetes and rates are expected to increase to 3.1 million by 2050.[1] Diabetes can lead to several significant complications including diabetic retinopathy, the leading cause of preventable blindness in adult Australians. There is a significant need to innovate in the prevention and treatment of diabetes to meet the increasing burden of disease.

Artificial intelligence is a broad concept that means enabling computers to act like humans, such as understanding human language, recognising images, and making decisions. Machine learning is a specific method within the field of artificial intelligence, where computers learn patterns from data and make decisions. Machine learning works by feeding a model large amounts of data so it can learn how to perform specific tasks. In healthcare, there is a wide range of data that can be fed into a machine learning model, such as electronic medical records, imaging, and data collected from an individual's wearable or phone. The best way to test a machine learning model is by using new, separate data from different locations with diverse patient populations, which helps ensure the model will work well when used in real clinical settings around the world. Deep learning is a subset of machine learning. The core difference is that in deep learning, there is no human guidance around what to look for in the data - deep learning models can figure out what is important on their own.

Given how rapidly artificial intelligence is changing healthcare, we conducted this evidence snapshot in order to explore how artificial intelligence/machine learning is being used in diabetes care.

Literature search

How did we answer this question?

This evidence snapshot uses a non-systematic approach, rapidly reviewing the most relevant, recent, and high-quality evidence to answer this question. The evidence was reviewed alongside one academic expert(OM) and one clinical expert (KC), to produce a brief evidence summary that is "good enough" to inform health services of relevant topics.[2] This document alone is not sufficient to solely inform decision-making.

The DELIVER research program is supported by a Commonwealth funded MRFF Rapid Applied Research Translation Grant

Findings

In the field of artificial intelligence/machine learning for diabetes prevention and treatment, there are two areas of rapid growth reviewed here: risk prediction and diabetic retinopathy.

Risk prediction

Traditional diabetes risk prediction models have significant limitations, including reliance on a limited number of risk factors and data sources, which fail to capture the complex interplay among biological systems involved in the development of the disease. Machine learning and deep learning are two methods of artificial intelligence that are capable of analysing very large biomedical datasets including electronic medical records, medical imaging, multi-omics data, and environmental data to identify high-risk individuals, identify novel risk factors, and ultimately build precision medicine responses.[3] In addition, machine learning and deep learning approaches can analyse multiple, very large datasets within a single model, termed multimodal models. Such an approach enables identification of complex interactions between various datasets which may not be apparent when using a single dataset.

To date, more than 40 different machine learning models have been developed and tested to predict diabetes risk.[3] A quarter of the studies were multimodal models, which consistently showed superior performance to single dataset models. While many of the single and multimodal models individually showed good performance (i.e., were able to predict diabetes risk with accuracy), there are several challenges:

- There are no standardised evaluation metrics that allow researchers to compare the performance of different models
- Almost no models are validated outside of the original dataset they are built on, including on diverse demographic samples
- Many models suffer from interpretability issues, in that it was challenging to understand how the model predicted diabetes, which has limited clinical utility
- Developing multimodal models is extremely time-consuming and as a result, it is challenging to quickly and easily scale such models.[4]

Diabetic retinopathy detection

For the past twenty years, much work has been done in developing machine learning and deep learning models for detection of diabetic retinopathy from fundus images, reflecting the significant advances in medicine more broadly around imaging-based artificial intelligence.[5] The increase in diabetes prevalence and ophthalmic workforce shortages, particularly in rural and regional settings, has created a need for screening tools outside these settings.

Of particular note, an Australian study of metro Victoria endocrinology services and remote Indigenous medical services, a deep-learning model from a Chinese dataset was tested, with very high performance accuracy across the sample of 96.9%.[6] When comparing performance of the model for detecting presence of diabetic retinopathy between metro and Indigenous samples, the model performed best on the metro sample, but was still as accurate as 85.9% for the Indigenous sample. Participant satisfaction and intention to use the screening service again was also very high.

What does this mean for health services?

Nearly all of the studies reviewed here did not use new, separate data, and instead tested their model on the same dataset it learned from, which can lead to excellent, but non-generalisable performance of the model. For Australia, the generalisability of AI models cannot be understated, as most training datasets come from the US or other countries. Australian samples differ in important ways. For example, Indigenous Australians, highly vulnerable to diabetes, are not included in international datasets, meaning the performance of these models on data from Indigenous Australians must be tested. For health services, the rapid and largescale adoption of artificial intelligence in healthcare cannot occur before such problems in data quality and bias are addressed, or else we risk entrenching health inequity.[4]

The DELIVER research program is supported by a Commonwealth funded MRFF Rapid Applied Research Translation Grant

While multimodal risk prediction models may be ideal from a disease standpoint, they still face significant challenges including time-consuming development processes, data merging difficulties, lower scalability, and complex interpretability issues, all of which can impede clinical adoption. A complex, richer prediction model that is uninterpretable clinically (i.e., clinicians cannot determine the underlying mechanisms behind prediction) is not useful for health services.

While numerous AI devices for diabetic retinopathy screening show promising performance and potential cost-effectiveness, significant knowledge gaps remain including the lack of head-to-head comparison studies to help clinicians choose which devices to deploy, and uncertainty about true implementation costs.

Limitations

- This review focused on just two growing areas of artificial intelligence in diabetes, risk prediction and retinopathy screening, and does not reflect the state of the evidence in other areas of diabetes.
- The literature base is expanding rapidly alongside the speed of technological change, and this summary may become out of date within a short time.

References

1. Zhou B, Rayner AW, Gregg EW, et al. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants. *The Lancet*. 2024;404(10467):2077-2093.
2. Norman G, Wilson P, Dumville J, Bower P, Cullum N. Rapid evidence synthesis to enable innovation and adoption in health and social care. *Systematic Reviews*. 2022;11(1):250.
3. Mohsen F, Al-Absi HR, Yousri NA, El Hajj N, Shah Z. A scoping review of artificial intelligence-based methods for diabetes risk prediction. *npj Digital Medicine*. 2023;6(1):197.
4. Wang SC, Nickel G, Venkatesh KP, Raza MM, Kvedar JC. AI-based diabetes care: risk prediction models and implementation concerns. *NPJ digital medicine*. 2024;7(1):36.
5. Rajesh AE, Davidson OQ, Lee CS, Lee AY. Artificial intelligence and diabetic retinopathy: AI framework, prospective studies, head-to-head validation, and cost-effectiveness. *Diabetes care*. 2023;46(10):1728-1739.
6. Scheetz J, Koca D, McGuinness M, et al. Real-world artificial intelligence-based opportunistic screening for diabetic retinopathy in endocrinology and indigenous healthcare settings in Australia. *Scientific Reports*. 2021;11(1):15808.

The DELIVER Research Project:

- Identifies what the people and healthcare providers of western Victoria need most in terms of home-based healthcare services
- Designs and tests the best way to deliver these services, so that home-based healthcare services will continue to grow and improve across the region and beyond
- Supports the growth of research in western Victoria, so that future research findings can quickly be translated to improvements in healthcare

---

August 2025

The DELIVER research program is supported by a Commonwealth funded MRFF Rapid Applied Research Translation Grant