CENTER FOR
THE GOVERNANCE
OF CHANGE

CASE STUDY 2: THE COLLABORATION OF DEEPMIND HEALTH WITH MOORFIELDS EYE HOSPITAL

Saif Ahmad, Cambridge University Hospitals.

Innovation, Sustainability, and the Future of Healthcare

In 2016, DeepMind Health began a collaboration with Moorfields Eye Hospital to assess whether deep learning techniques could be used to interpret optical coherence tomography scans (OCT) to help triage and diagnose the health of the retina in patients with visual problems. A landmark research paper was published in 2018 demonstrating that the deep learning method produces results comparable to those of experts in the field, and that this was generalizable across imaging devices. Importantly, DeepMind Health are focused on making it possible for clinicians to understand how the algorithm arrived at its conclusion [1]. DeepMind Health have also been mindful to promote a positive perception amongst clinicians and the wider public. How they have attempted to do this, and its effectiveness is the focus of this case study.

BACKGROUND

DeepMind Technologies Limited was founded in September 2010, by artificial intelligence (AI) experts Demis Hassabis, Shane Legg and entrepreneur Mustafa Suleyman. Their vision, as stated on their website, was that they "wanted to try a new interdisciplinary approach to accelerate the field – bringing together new ideas and advances in machine learning, neuroscience, engineering, mathematics, simulation and computing infrastructure – as well as a new way of organising scientific endeavour".

DeepMind's early research and development demonstrated that deep reinforcement learning, a type of machine learning based on cumulative rewards, could be used to successfully learn control policies from high-dimensional sensory data. On a more tangible level, what they showed in their research in 2013 was that their technology enabled computers to learn the control policy of, i.e. how to play, seven different Atari computer games from scratch [2].

This and other breakthroughs attracted Google to purchase the company for $400m in January 2014, reportedly following a failed acquisition by Facebook in December 2012 [3]. Importantly, as part of the DeepMind deal Google announced that they would be setting up an independent AI ethics board to ensure the technology was not abused.

In February 2015 a seminal paper published by researchers at DeepMind in Nature described how their technology enabled one single algorithm to learn 49 different Atari games essentially from scratch (i.e. only knowing that the input data were visual images, and the number of actions available in each game). The algorithm was able to play the games at skill levels comparable to a professional games tester [4]. The research abstract states that: "This work bridges the divide between high-dimensional sensory inputs and actions, resulting in the first artificial agent that is capable of learning to excel at a diverse array of challenging tasks." This paper has been cited over 2000 times since its publication, demonstrating the significance of this research. The findings had clear implications across numerous skill domains and the potential for this technology to be applied to healthcare was soon realized

–by 2016, leveraging new collaborations with healthcare providers and researchers, DeepMind Health was formed.

One of the first and most successful collaborations for DeepMind Health was with Moorfields Eye Hospital and this case study will focus on its outputs and use the collaboration as an exemplar to demonstrate how DeepMind Health strived to ensure that the applications of its deep learning technology for healthcare had the best chance to be adopted in the wider clinical arena.

DEEP LEARNING APPROACHES TO IMPROVE DIAGNOSIS OF RETINAL DISEASES

In June 2016, Moorfield's Eye Hospital in London and DeepMind Health announced a new collaboration to explore how AI could be applied to evaluating eye scans to detect signs of disease more quickly and accurately than medical professionals. The press release from that time includes quotes supporting the collaboration from three separate eye health charities [5].

In just over 2 years, this collaboration had led to a landmark paper being published in Nature Medicine which described how machine-learning techniques could be used to evaluate optical coherence tomography (OCT) scans to triage and diagnosis patients with visual symptoms [1].

OCT scans provide three-dimensional information of a patient's retina and are routinely used by clinicians to help assess patients who present with symptoms such as visual loss. In this study, the OCT data was transformed into a three-dimensional map, which could be interpreted using a machine-learning algorithm which then classified each map based on the probability of it reflecting different diagnoses. This allowed the algorithm to triage patients as to how urgently they needed to be seen.

To develop the automated methodology to convert OCTs into three-dimensional maps, 877 training images were manually segmented by ophthalmologists to represent different tissue types (for example, different parts of the retina). Each image consists of 128 individual slices, however only three to five representative slices needed to be manually segmented for each image. Once the tissue segmentation network was created, the classification network was generated using a training set of 14,884 OCTs which had a confirmed diagnosis and referral decision (see Figure 1).

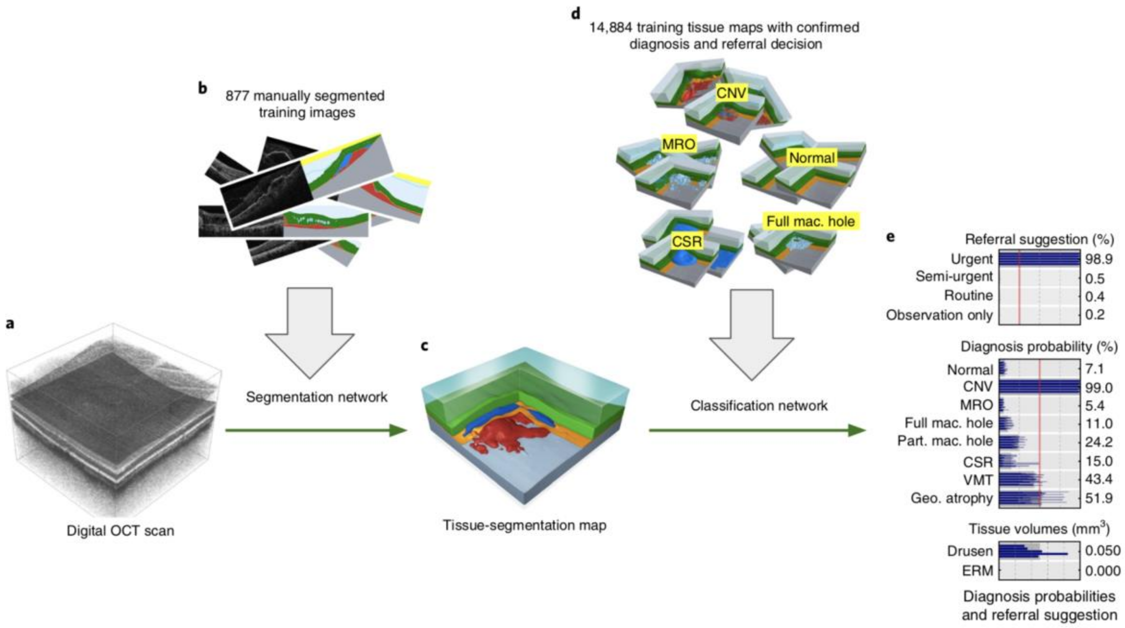

Fig. 1 | Our proposed AI framework. a, Raw retinal OCT scan ($6 \times 6 \times 2.3\,mm^3$ around the macula). **b**, Deep segmentation network, trained with manually segmented OCT scans. **c**, Resulting tissue segmentation map. **d**, Deep classification network, trained with tissue maps with confirmed diagnoses and optimal referral decisions. **e**, Predicted diagnosis probabilities and referral suggestions.

Figure 1. Description of workflow to assess retinal photography [1]

These steps developed the framework which was then tested on 997 patients' cases – these were not included within the training set. This analysis showed that the error-rate of referral decisions for this AI-based framework was only 5.5% which was comparable to clinical experts.

Importantly, both the segmentation and classification networks were trained on five separate instances each. This allowed five separate hypotheses to be developed for each scan which could be used to inform clinicians of areas of ambiguity. Specifically, some areas within the OCT may be difficult to assign to a distinct tissue type. These areas could be highlighted, and it would thus be possible for clinicians to look at these areas themselves to understand what was underlying the ambiguity.

Moreover, the researchers showed that their approach could be modified to interpret OCT images from a different OCT device. The segmentation step was repeated for this new scanner, i.e. where OCT images were initially manually segmented. However, the classification step remained the same as that used from the initial set of experiments, and this led to similarly low error-rates in triaging patients. This workflow offered clear advantages in terms of applying it to the clinic, as the more resource-intensive step of training the classifier is not required for each device.

The researchers stressed in their conclusion that this approach has the potential to be generalized across a number of different medical imaging-based applications. For instance, DeepMind Health is currently exploring whether similar approaches could be applied to mammograms as part of breast cancer screening [6].

In 2019, DeepMind demonstrated a prototype device based on this technology which they hope to take to regulatory approval. In the demonstration, a patient was scanned and their images were interpreted using a cloud-based platform, with a referral suggestion made in 30 seconds [7]. It is likely that a clinical trial will be performed in due course to evaluate the technology's effectiveness, and health economic benefits in a clinical setting. Of note, IDx, a US-based company received FDA approval for a device that uses an AI-based approach to diagnose diabetic retinopathy, after a clinical study in 900 patients [8].

STRATEGIES THAT DEEPMIND HEALTH HAS EMPLOYED TO REDUCE BARRIERS TO ADOPTION OF AN AI-BASED HEALTH TECHNOLOGY

This case study demonstrates several strategies that have been employed to shape the perception of the technology being developed. The following sections summarize these strategies, and how they are important for various different stakeholders (see Figure 2).

Ensuring accuracy and feasibility

Arguably the key determinant influencing the likelihood of this technology reaching the clinic is how well it works. The data demonstrates how the error-rate of their method is comparable to experts in the field. The technology could also be adapted relatively easily, to incorporate OCT data from different devices. Moreover, at the publicity event linked to the prototype device, DeepMind were able to show that a decision could be arrived at within 30 seconds.

Ethics, data privacy and trust

Significantly, DeepMind set up an Independent Reviewers Panel (IRP) at the time of creating DeepMind Health. This panel consisted of 9 members including key opinion leaders in health policy and former MPs. To ensure independence, members were not paid, bound by any secrecy agreements and were free to engage with any external professionals as they saw fit. The IRP were given a budget (around £50,000) to commission work and fund an administrative team. The IRP's objective was to hold DeepMind Health to account in order to operate to high ethical standards. The panel's only specific responsibility was to produce an annual report and it published two reports (in 2017 and 2018) prior to being disbanded in 2019 after DeepMind Health was incorporated into Google Health [9], [10].

Issue	Patient	Public	Healthcare professional	Regulator
Accuracy of device	3	2	3	3
Feasibility of device	3	1	3	2
Data protection	3	3	3	3
Quality control	2	1	3	3
Understanding of how algorithm works	2	1	3	3
Potential to use algorithm with other OCT devices	1	1	3	1

Figure 2. How specific issues are viewed by relevant stakeholders (1-3, where 1 is not important and 3 is very important)

The 2017 report recommended increasing patient-public involvement and engagement with medical royal colleges in order to better understand complex healthcare needs, and evaluated DeepMind Health's data privacy measures. The 2018 report reviewed the progress made following the first report and also recommended that DeepMind Health clarify its relationship with Google, be anti-monopoly, show greater transparency of their proposed revenue models and seek to only make reasonable profits from its dealings with the public sector.

Data privacy is a critical issue in the context of digitizing health records and information. Although the Moorfield's Eye Hospital collaboration has not attracted any specific concerns, in July 2017 the Information Commissioners Office (ICO) ruled that a subsequent collaboration between Royal Free Hospital and DeepMind had failed to comply with the Data Protection Act when data from 1.6 million patients was transferred as part of a kidney disease trial [11]. There were particular concerns about the transparency of the agreements between the hospital and DeepMind [12]. This provided significant negative publicity for DeepMind Health although most of the actions recommended by the ICO were the responsibility of the Health Trust rather than DeepMind.

In terms of communications related to Moorfield's collaboration, a Patient Engagement Event was held in September 2016 (shortly after the collaboration was announced) and was attended by over 130 members of the public [13]. A key part of the agenda was a session on implications of sharing patient data. It is also notable that ongoing press releases include a number of independent charities that are approached for quotes, helping to emphasize that the non-commercial value of this research.

Quality control and the "black box problem"

A key concern around AI-based health technologies is that externally it may not be possible to understand how an algorithm has arrived at its decision. The workflow within DeepMind's framework allows demonstration of areas of ambiguity within the OCT image. The authors cite this as key in understanding why a decision was arrived at and claim that it will also facilitate quality control measures. Solving the "black box problem" is of vital importance in enabling clinicians, regulators and patients to trust an algorithm [14].

DEEPMIND HEALTH AND ITS RELATIONSHIP WITH GOOGLE

Although Google had bought out DeepMind in 2014, part of the takeover had allowed DeepMind to remain independent of Google and ensured that DeepMind Health operated independently of Google's health enterprises (Google Health).

However, in late 2018 it was announced that DeepMind Health would be subsumed into Google Health. This process was completed in September 2019. Notably as part of the move, the IRP was disbanded as this UK-centric panel was not deemed suitable for the international market that Google Health seeks to operate in. Concerns have been raised by many about data privacy, given that DeepMind Health partnered with NHS Trusts with the understanding that the organization was distinct from Google [15]. As of now, Moorfield's has decided to continue its collaboration and now has an agreement in place with Google. Indeed, five NHS trusts agreed to carry over the DeepMind Health agreement with Google and only Yeovil NHS Trust declined to do so. Yeovil NHS Trust cited the reason for this being that they did not intend to operationalize the application (streams) that the agreement was for, rather than any privacy concerns around Google [16].

CONCLUSIONS

This case study highlights an exciting intersection of AI-based technology with an important health need. Medical imaging data is in many respects a relatively low-hanging fruit in terms of applying AI to health and this example illustrates some of the opportunities and challenges when applying AI to medical imaging data. It will be intriguing to see how pathways in retinal imaging change over the coming years as AI seeks to play an increasing role. How effectively this transition takes place may well reflect how efficiently AI has a positive impact on healthcare in general.

REFERENCES

[1] J. De Fauw *et al.*, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nat. Med.*, vol. 24, no. 9, pp. 1342–1350, Sep. 2018.

[2] V. Mnih *et al.*, "Playing Atari with Deep Reinforcement Learning," Dec. 2013.

[3] Forbes, "Google To Acquire Artificial Intelligence Company DeepMind." [Online]. Available: https://www.forbes.com/sites/amitchowdhry/2014/01/27/google-to-acquire-artificial-intelligence-company-deepmind/#7e00abd715b5. [Accessed: 15-Mar-2020].

[4] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[5] "Excited to announce a new medical research partnership with DeepMind Health," Moorfields Eye Hospital NHS Foundation Trust [Online]. Available: https://www.moorfields.nhs.uk/content/excited-announce-new-medical-research-partnership-deepmind-health. [Accessed: 15-Mar-2020].

[6] S. M. McKinney *et al.*, "International evaluation of an AI system for breast cancer screening," *Nature*, vol. 577, no. 7788, pp. 89–94, Jan. 2020.

[7] M. Murgia, "Artificial intelligence group DeepMind readies first commercial product," Financial Times, 31-Mar-2019. [Online]. Available: https://www.ft.com/content/0e099914-514a-11e9-9c76-bf4a0ce37d49. [Accessed: 15-Mar-2020].

[8] FDA, "FDA permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems." [Online]. Available: https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence-based-device-detect-certain-diabetes-related-eye. [Accessed: 15-Mar-2020].

[9] M. Bracken *et al.*, "DeepMind Health Independent Review Panel Annual Report," Jul. 2017.

[10] M. Bromiley *et al.*, "DeepMind Health Independent Review Panel Annual Report," Jun. 2018.

[11] ICO, “Royal Free - Google DeepMind trial failed to comply with data protection law.” [Online]. Available: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2017/07/royal-free-google-deepmind-trial-failed-to-comply-with-data-protection-law/. [Accessed: 15-Mar-2020].

[12] J. Powles and H. Hodson, “Google DeepMind and healthcare in an age of algorithms,” *Health Technol. (Berl).*, vol. 7, no. 4, pp. 351–367, Dec. 2017.

[13] “DeepMind Health patient engagement event," Moorfields Eye Hospital NHS Foundation Trust [Online]. Available: https://www.moorfields.nhs.uk/content/deepmind-health-patient-engagement-event. [Accessed: 15-Mar-2020].

[14] A. J. London, “Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability,” *Hastings Cent. Rep.*, vol. 49, no. 1, pp. 15–21, Jan. 2019.

[15] H. Chowdury, “Google takes control of DeepMind Health despite privacy concerns.” [Online]. Available: https://www.telegraph.co.uk/technology/2019/09/18/google-takes-control-deepmind-health-despite-privacy-concerns/. [Accessed: 15-Mar-2020].

[16] Financial Times, “NHS trusts sign first deals with Google.” [Online]. Available: https://www.ft.com/content/641e0d84-da21-11e9-8f9b-77216ebe1f17. [Accessed: 15-Mar-2020].