The future of healthcare algorithms in Aotearoa New Zealand

Attribution
Orion Health Intelligence wishes to thank all research participants for their generous contributions to this project. Special thanks to Rebecca Caroe (specialist copywriter and B2B marketer) and researcher Dr Karol Czuba (Managing Partner - WeCare Health) and the project advisors Professor Alain Vandal (University of Auckland), Haze White (Tāmaki Regeneration Company) and Precision Driven Health. Researchers from Orion Health Intelligence were Pieta Brown (Principal Investigator) and Rachel Owens.

The Algorithm Scan project

Precision Driven Health is an award-winning research partnership between New Zealand's health IT sector, health providers, and universities to improve health outcomes through data science. In the Algorithm Scan project, we sought to describe the New Zealand healthcare algorithm landscape, focusing on predictive algorithms for decision support. This study involved in-depth interviews with 35 health sector representatives, a literature review and an online survey.

This decade, the overarching vision of a learning, adaptive health system is within our grasp, and predictive models will likely provide important insights that underpin this change. By understanding the current state, opportunities and risks, we can frame a pathway towards successful adoption.

Data sharing, modelling, and digital innovations have characterised our response to COVID-19. Building on this activity, now is the time for us to plan the future for healthcare algorithms in Aotearoa New Zealand.

Findings

Current state	Aotearoa New Zealand has a well-governed, ethical, and relatively uniform health system, a values-driven clinical workforce, enthusiastic technology and research workforces, and rich health data. We are well positioned to safely and effectively integrate predictive models into healthcare. Clinicians and administrators already use over 50 algorithms and models; however, wider adoption requires clean and timely data, skills and funding.
Opportunities	Opportunities for algorithm use in healthcare include prediction in clinical practice, operational efficiencies, the development of data science literacy and expanding the space for empathy in the doctor-patient relationship. However, we must build trust in models to take advantage of these opportunities. Shared knowledge, model transparency and provision for model feedback will help us build that trust and overcome concerns about reliability.
Risks	As we adopt models, we must manage risks such as potential bias in model outputs, inequity, use of the correct data, ethics, transparency, clinician hesitancy, and liability concerns.
A diverse foundation with a local focus	We must operate on a foundation of high quality and accessible data and diverse, multi-disciplinary teams. Models should be evaluated, calibrated, and developed locally within te ao Māori context.
Governance	The dearth of inventory, reporting or monitoring of health algorithms poses a risk to our healthcare system. We could mitigate this risk with national standardisation and oversight, and sharing best practices concerning algorithm use. Risk oversight could encompass the preservation of clinical intuition.
Productive workflows	Algorithms should weave seamlessly into clinical workflows, enhancing productivity with defined model options, and appraisal and impact assessment feedback opportunities.

Algorithms in healthcare

Clinicians draw upon health data from diverse sources when making diagnosis and treatment decisions. Over time, that data has grown in quantity, relevance and intersectionality, posing a challenge for clinical practice. It is increasingly possible to support or even automate healthcare decisions by improving data collection, processing, and reasoning to meet that challenge. These data calculations and interpretations involve the use of algorithms.

In general terms, algorithms are sets of rules or processes for making calculations or understanding problems, including predictive models. The Algorithm Charter for Aotearoa New Zealand defines predictive models as:

> **"…models which make predictions about some unknown variable, based on one or more known variables."**

These models range from simple rules-based calculators to models built using sophisticated methods such as neural networks.

Using patient data, clinicians can employ predictive models to support accuracy and efficiency in decision-making for better health outcomes and reduced time and cost of care. These models process what is humanly 'unprocessable' due to complexity, confounding, and bias.

As clinicians increasingly rely on algorithms in varied contexts, we must carefully consider algorithm selection, design, development, deployment and monitoring. We believe there remains significant potential for algorithmic tools to change our health work practices and improve healthcare outcomes. Aotearoa New Zealand is well positioned to adopt these tools safely and effectively.

There is enthusiasm for this change within our research and technology sectors. However, we are often hampered by a lack of access to clean and timely data, skills and funding. Our healthcare sector has the opportunity to leverage local and international innovation and avoid the pitfalls that could damage clinician and consumer trust.

Current state

Our research and interviews identified over 50 algorithms and models in use within the Aotearoa New Zealand healthcare sector, the majority of these within hospital settings. These algorithms and models support clinical decision-making across prognosis, diagnosis and treatment.

Interviewees described algorithms and models that estimated the likelihood of outcomes such as developing deep vein thrombosis, cardiovascular disease events, complications of procedures, and benefit from an intervention. Clinicians also used models to guide treatment decisions by informing drug dosage or cancer therapy options. Managers and planners used operationally-focused models to predict the length of hospital stays to support planning and rostering activities.

Example

Prognostic models estimate the risk of a future outcome for a patient by using information about an individual such as their age and comorbidities. One such model is nzRISK. This model supports shared decision-making between patients and specialists by providing patient-specific mortality risk estimates for a surgical procedure.

Algorithms in District Health Boards

At most District Health Boards (DHBs), algorithms are accessed via existing web-based tools, coded in-house by the DHB or integrated with third-party software. A small number of clinicians or technical staff may maintain versions coded in-house.

In most cases, health providers do not maintain an inventory of algorithms and models in use. Neither do they have the capacity for reporting and monitoring the use and performance of those algorithms over time.

Software coded locally comes under the risk assessment, due diligence, and governance oversight of the local DHB, and the current implementation process carries risk at each step.

Errors in algorithms may risk systematic impairment to our healthcare system. There is an opportunity for national standardisation and sharing best practices; Aotearoa New Zealand does not yet have a trusted library of nationally approved algorithms nor agreed clinical practice and record-keeping of algorithm use and application.

Algorithms in clinical practice

Our clinical participants represented a range of DHBs, primary health organisations and research institutions. Many of the clinicians interviewed expressed awareness and optimism around the opportunity for the increased use of algorithms and models in clinical practice.

When asked about the most significant value algorithms will bring over the next five years, 47% of our survey respondents said improved quality of care.

Clinicians are already using algorithms in their practice. In the past four weeks, one-third of survey respondents had used a model over 15 times, and 30% between six and 15 times.

The number of algorithms and frequency of use for a clinician depends on their medical specialty and personal preferences. While most clinicians use a small number of models, these are not consistent nor streamlined into working practices, and the use of more advanced tools/models is limited. Of the models used, 64% estimated risks to inform clinical decision-making, and 35% supported treatment choices. The sample is small (with the risk of respondent bias); however, these results are consistent with regular and sustained algorithm use.

Clinicians access algorithms and models from various places, including apps, external websites, and in-house software applications. Senior clinicians mainly guide the use of a tool by junior staff within their locality and specialty.

Yet, many clinicians still reserved judgement about algorithms in healthcare. Dr Doug Campbell is the co-developer of the nzRISK model. He described how, in his view, hesitancy is often because there is a perception that these models are unreliable.

“

For many clinicians there’s a ‘great suspicion’ of algorithms.

Dr Doug Campbell,
Anaesthetist at Auckland DHB

However, he remains optimistic about progress within the next five to 10 years. He believes we will see a democratisation of algorithm use over this period. Dr Campbell suggests this is particularly likely when a few well-chosen risk tools in a specialty can assist with decision-making and deliver quality information to patients and clinicians.

Algorithmic opportunity

Opportunities involving healthcare algorithms point to an exciting future for healthcare in Aotearoa New Zealand.

Our relatively uniform health system, balanced and sensible ethics committees, and universal National Health Index (NHI) provide a robust platform for increasing the adoption of algorithms in clinical practice.

> We have better data than most countries (data from the health and social systems). Few countries can link these in the way we do.
>
> **Professor,**
> University of Auckland

Opportunities include clinical prediction, operational efficiency, building data science literacy, and expanding the space for empathy in clinical practice.

Clinical prediction & operational efficiency

From the clinical perspective, opportunities range from tailoring treatments with genomic information to imaging use-cases where algorithms already support radiology. These opportunities can also address operational pressures, such as the high cost of manual image interpretation. If the workforce becomes constrained, radiologists could focus their skills on complex imaging and diagnostics with algorithmic support for other tasks.

Data science literacy

Another opportunity is building data science literacy within the healthcare sector. Several interview participants described examples of this opportunity within medical school training. We could advance the acceptance of models by leapfrogging our educational efforts directly into the newest cohort of clinicians in training. We could integrate training in safe algorithm usage through the early years of their careers.

Empathy

Professor Tim Dare from the University of Auckland described how well-designed and reliable tools could reinforce the quality of the doctor-patient relationship by creating space for more human aspects. We know empathy and compassion enhance the quality of that relationship. The suggestion that big data will destroy or undermine this is scaremongering.

> As long as you can see that it's safe and effective. [I believe] the role of the doctor will change and adapt.
>
> **Professor Tim Dare,**
> University of Auckland

Building trust

Building trust, often through explainability, was a recurring theme in our research interviews. Clinicians generally agreed that they don't need to know the nuances of how a model works. Clinicians did, however, need to know enough to use and trust an algorithm.

Trust through shared knowledge	Many interviewees suggested we haven't met a minimum educational requirement and that the current approach is too technology-centric. Dr Matthew Strother cited this as a critical challenge for algorithm adoption: "Have we provided sufficient education for the average clinician?" To encourage adoption, we must therefore draw on deep knowledge of the practice of clinical decision-making, the human factors involved, and the art and science of medicine. ***Data science adoption requires us to trust not only in the expertise of new roles but also in the data and technology supporting those roles, which can create tensions.*** Roles such as 'data scientist' and 'machine learning engineer' are relatively new and not typically part of established clinical teams and informatics career pathways. Karen Day is a Senior Lecturer at the University of Auckland, health informatician and trained nurse and midwife. She described how these new roles are challenging for clinicians who share knowledge and expertise and clearly understand how clinical roles work together.
Trust through transparency	***To build trust, clinicians said they need visibility and timely access to model risks and limitations.*** Model documentation should include links to academic papers and peer-reviewed assessments. Participants described how they wanted clarity on the caveats that apply when using a model. They do not have time to dig through the finer details of published papers, let alone validate their conclusions with an expert statistician.

Trust through feedback

One way to enhance trust and acceptance could involve building human validation mechanisms into software that delivers model-based decision support.

Dr Yaniv Gal, Chief Technology Officer of Kāhu – a spinoff company of MoleMap – successfully applies AI to thousands of images to support early skin cancer detection. Discussing how to build trust in healthcare data science, he said that “the way to gain trust in algorithms is by experimentally validating their performance.”

Modelling software could prompt clinicians for feedback such as, “do you agree with this result?” or “did you find this result useful?” This approach may also provide software developers with helpful information for improving implementation and user experience by uncovering user issues and pain points.

Risks

Assessing algorithmic risk is a rapidly evolving skill set. We must address known risks; however, the faster we collect data and develop new tools, the faster we will uncover previously unknown risks requiring resolution.

Bias and inequity

We heard concerns about worsening health inequity through biased models and data, balanced against a significant opportunity for algorithms and models to help monitor and mitigate bias.

Models could worsen inequity if historical biases are programmed in and scaled through the deployment of algorithms and models.

Karen Blake is Director at PwC New Zealand and previously Head of Clinical Informatics at Health Alliance. She believes we must focus on diseases of inequity and those associated with poverty, such as diabetes, chronic obstructive pulmonary disease, respiratory disorders and childhood obesity. She urges us to take a considered and cautious approach.

> We need to be really careful that we don't increase inequity in the system through using AI.
>
> **Karen Blake,**
> Director PwC NZ

Interviewees also raised concerns that emerging data science efforts, such as direct-to-consumer genetic test results, will reflect the 'worried well'. Our efforts might only improve the health of the most well off.

The right data

Clinicians agreed that we need the right data in place to support safe and effective algorithms and models. They described how important sources are often missing from the data available for analysis. Interviewees identified data gaps affecting patient outcomes, including social determinants of health and healthcare system factors such as hospital capacity, staffing, time of the day and day of the week.

Ethical AI

Many participants described concerns about how ethical guidelines can be applied as a box-ticking exercise to 'fudge' and obfuscate issues. The ethical challenges are often subtle and nuanced. The slow and difficult work to address these challenges was deemed unattractive to commercial organisations and researchers.

Hidden complications

Several clinicians described specific concerns about the 'devil in the detail' for predictive model performance metrics. They described the difficult work of digging into the detail of a published model to critically evaluate it for use in the context of the 'nitty gritty' of what happens at the bedside.

Hesitancy and liability

Hesitancy around data science and the underlying data quality is another challenge. Rochelle Style, AI governance and ethics consultant, explained that it's about finding an appropriate balance between risk and benefit.

> I don't think we should wait until everything is perfect because then we'll never do anything. But equally, we don't want to let the genie out of the bottle until we've done appropriate due diligence.
>
> **Rochelle Style,**
> Consultant

Most clinicians also raised concerns about recourse and liability if something does go wrong based on the implementation or use of a decision support tool.

A pathway towards adoption

Our study suggests five core elements that would support a framework for meaningful adoption and use of algorithms: Getting the foundations in place, Local validation and te ao Māori, Governance and risk oversight, and Workflow integration.

Getting the foundations in place

Many clinicians cited pressure to adopt data science in healthcare, referring to the success achieved in other industries as a contributing factor. A cautious approach was advocated, "We need the right structures in place first," says Dr Alex Kazemi. A foundation for algorithm adoption in healthcare involves high-quality data and bridging the gap between clinical and technical disciplines.

High-quality data

> ***We will realise the clinical potential for algorithms from a foundation of high-quality, accessible data.***

The UK government has invested heavily in health data research and informatics; however, it took the COVID-19 pandemic for much of that work to be taken seriously and used. Dr Ben Goldacre and the Oxford DataLab published "Factors associated with COVID-19 death in 17 million patients", a paper based on open trials, prescribing and a large base of pseudonymised patient records. This rapid response analysis has informed public health policy in the UK since July 2020, highlighting the value of quality data in healthcare and health management.

> "
>
> COVID-19 shows more clearly than ever that we can and must deliver clean, real-time, standardised data to support direct care and all aspects of system planning and response.
>
> **Dr Ben Goldacre,**
> Director, Oxford DataLab

While we have rich data sets, there are often significant challenges in accessing high-quality and timely data to support decision-making. As Professor Colin Simpson of Victoria University explained, "[a] quick win is really just providing clinicians with information and allowing them to act upon it."

Collating and cleaning data for quality are exacting yet essential tasks. While the importance of quality data is understood, investment in data quality initiatives is often limited. Data scientists and researchers may gravitate towards developing new models over this 'unseen' foundational work.

Bridging clinical and technical disciplines will be vital in creating a framework of acceptable operational norms around algorithm design, use and incorporation into clinical practice.

A Professor of Health Economics explained that success in hard problems comes from multi-disciplinary teams using design thinking methods. Dr Alex Kazemi, a specialist doctor and writer, envisions that this skill base might broaden to include data visualisation experts, social scientists and ethicists. His experience through decades of clinical practice is that people respond to narratives and storytelling supported by practical examples. For strongly values-driven clinicians, these stories need to convey clearly how their patients and colleagues will benefit.

Style endorses the need for both multi-disciplinary and diverse teams.

> Every issue seems to have so many layers - ethically, culturally, scientifically and politically.
>
> **Rochelle Style,**
> Consultant

We should consider how clinicians incorporate predictive model outputs (such as a numeric score) into diagnostic and treatment decisions. A clinician may slot that number into their diagnosis or review it to confirm their judgment-based diagnosis. However, if the number deviates from the clinician's preconceptions, would they trust the algorithm less? And would that experience jeopardise model adoption?

From the University of Auckland, Professor Jim Warren and doctoral candidate Mike Merry emphasised the importance of understanding how a model will inform decisions before any technical development. As they explained, "You can't talk about model performance until you know the decision it's going to be used for. And, given that, you probably can't train the algorithm optimally unless you know what you're going to use it for."

Warren and Merry advise that we focus on the users and context for algorithms: a model's place in the workflow, the actionable information, sensible performance metrics, and clinicians' decision thresholds in practice.

Local validation

Local validation involves evaluating internationally developed models and building new models that are appropriate for our local communities. Clinicians described a difference between models that are 'physiological', modelling dynamics within a body system, which may have international applicability, and those that predict outcomes based on a broader range of factors about a person, which may reflect local health settings, cultures and demographics.

Humans train algorithms for specific tasks. Models may therefore incorporate cultural norms and values via data selection, methodology, and target outcome definition.

Many of the algorithms offered for use in Aotearoa New Zealand, were developed overseas based on populations with different characteristics from our local population and in other healthcare settings.

Of the 52 tools identified by our literature review, only 30 had been validated through research for local use. Of these 30, researchers found that some were unreliable, despite validation.

When considering the applicability of algorithms developed overseas, Professor Rod Jackson found that risk equations often don't include social determinants of health, such as poverty measures and access to healthcare.

> "
>
> Age, ethnicity, and deprivation are measuring a lifetime of exposure and dealing with things. [These are] surprisingly important predictors.
>
> **Professor Rod Jackson,**
> University of Auckland

Professor Jackson adds, "most commentators don't appreciate that the data you need to truly validate an overseas algorithm is the same data you need to develop a new local equation."

Style articulated the ethical issues around local validation. She explains, "[Algorithms] can have such a large impact on thousands of patients. You need to very carefully ask yourself - this algorithm has been trained on data from England (for example); is it beneficial and respectful, in an ethical sense, to use this with New Zealanders? This seems really fundamental to me."

Given that we often require the same data for model evaluation and development, there is an opportunity to adjust, recalibrate or develop local models for use in Aotearoa New Zealand. At a minimum, we may consider retrospectively validating models developed in other healthcare systems and populations.

Validating for groups

The inclusion of ethnicity as a factor in predictive algorithms raised questions and concerns.

Interviewees agreed that algorithms should perform well, be statistically robust, support action, improve health outcomes, and, importantly, not embed or perpetuate biases.

Diverse and inclusive teams will be critically important for developing safe and beneficial algorithms that will serve populations.

We already know that researchers have had difficulty reproducing results across hospital sites. So we should also establish model performance indicators for subgroups and share that information with governance groups. Further, we will need management processes to underpin model assessment, ethical review and localisation.

Te ao Māori

In Aotearoa New Zealand – where Māori are a Crown partner under Te Tiriti o Waitangi and have worse experiences and outcomes when engaging with the health system – bringing te ao Māori perspective to algorithms in healthcare is fundamental.

Te ao Māori context for algorithm development resonates with Dr Daniel Wilson, a teaching fellow at the University of Auckland, "I'm excited about this idea of having more than one worldview or value set inform the construction of algorithms."

Dr Wilson explained how developing safer systems via te ao Māori could serve us in Aotearoa New Zealand, and be world-leading:

"I'm optimistic that there will be better outcomes from taking in te ao Māori worldview, the interests of Māori, the timescales involved, [the] value concepts and considering algorithms in that frame as well as having more 'standard' values and metrics."

Professor Dare is similarly optimistic about delivering on this vision and incorporating multiple value sets in algorithm development. "One way is to build these in from the beginning and be guided by people who really know and understand these values and can say this application was built in a respectful and appropriate way and reflects the values in a way we recognise," he said.

Governance and risk oversight

In 2020 the Medical Council of New Zealand hosted a discussion paper focused on situations when artificial intelligence is involved in the care of patients. While its findings have not yet been made public, this is likely to play an important role in shaping our future clinical governance standards for algorithms.

The governance challenge

Style cited concerns around the lack of evidence of efficacy and auditing of algorithms in use. “There is currently a dearth of governance and robust frameworks around algorithm incorporation and validation in practice,” she said.

A surgeon at Canterbury District Health Board described the lack of confidence and the many unknowns around the adoption of algorithms. They see “[a] lack of governance [and] really knowledgeable people who can comment on risks and benefits. Everyone is a bit scared. Also scared of vendor ownership.” Researchers and entrepreneurs developing the tools which clinicians may want to use will face this challenge.

Governance teams will need an audit trail for verification and historic checking without overstepping privacy boundaries.

> ***There is a complex interplay between software ownership, securing proprietary data outputs, and sufficient transparency of model-based decisions for independent validation. No one has yet worked out how to cross this Rubicon with confidence.***

A standardised approach

Every organisation has an ethical view, assessment criteria and processes around health data, which we should standardise for consistency and transparency, according to a researcher at Te Pūnaha Matatini.[1]

> “
>
> The technical capability of our data system is rapidly outstripping our ethical and validation purview. [I] would be really keen on a single ethical review system for the use of any linked health data.
>
> **Researcher,**
> Te Pūnaha Matatini

Concerning health research data ethics, standardised processes may fall within the purview of the Health Research Council. The Council advises the Minister of Health on ethical issues in health research as part of its statutory role.

Preserving clinical intuition

Any new tool takes time to become commonly used. Clinicians vary in years of practice and digital literacy levels. Early-career clinicians in the digital native workforce will likely use algorithms more frequently.

Expertise in your field gives a sense of what is probably right when reviewing algorithm results. However, less experienced clinicians have not yet developed a depth of case histories in our unique population with particular health challenges. We should consider whether reliance on algorithms would impact the development of clinical intuition. Our educators and the tertiary institutions who guide early career training could address this challenge.

[1] Te Pūnaha Matatini is a Centre of Research Excellence focused on solving complex problems, hosted by the University of Auckland

Workflow integration

What would it take to make the most of algorithms in a clinical setting?

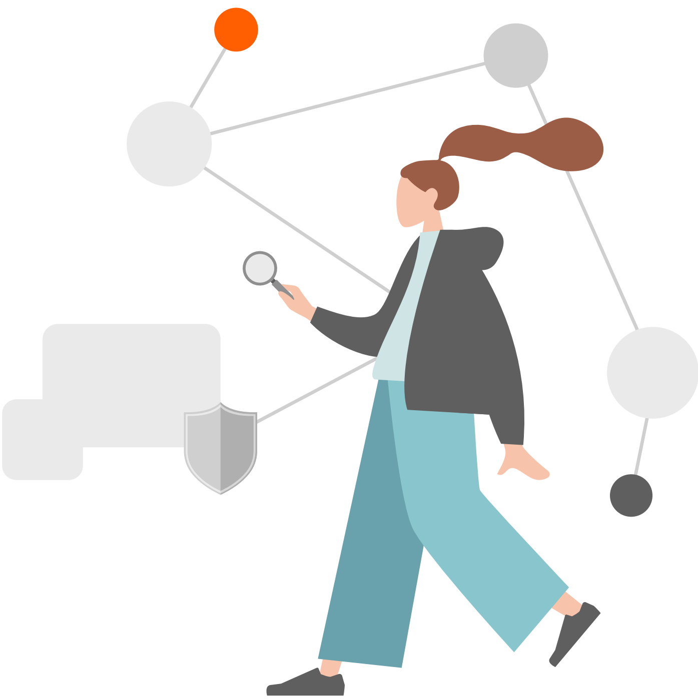
There is frequently more than one algorithm available for any given situation. Clinicians need access to an approved, structured process for selecting which to use.

A suite of nationally approved tools could simplify algorithmic choice.

Downstream, clinical workflows incorporating algorithm use will involve appraisal and impact assessment feedback as well as implementation. We do not want to hinder productivity for those who are already overworked, especially when the stated goal of many of these tools is to save time. Professor Jackson reflected on 30 years of effort supporting the adoption and use of cardiovascular risk assessment:

> Make it as easy, quick, and value-add as possible. Make the right thing the easy thing to do.
>
> **Professor Rod Jackson,**
> University of Auckland

We know that pragmatism wins out in healthcare innovation. So for workflow implementation, the clinician should be able to answer positively that the algorithm fits into their processes at the right moment, with information that's sufficiently valued and actionable. Each frontline clinician should be able to point out – without hesitation – why an algorithm is worthwhile for them.



Conclusion

Diverse teams are critical for developing and testing algorithms that suit the health services and population of Aotearoa New Zealand. Expert oversight of clinical algorithms is a specialist area – we are unlikely to find sufficient expertise in Aotearoa New Zealand within each healthcare locality. Therefore, our research suggests exploring solutions built around nationally centralised oversight.

It is still early days for healthcare algorithms. We know the strengths of our healthcare sector (values-driven workforce, ethics and governance) and the significant opportunities (localisation and collaborative problem-solving). We also know that it will take multiple iterations before we satisfactorily embed healthcare algorithms in our healthcare sector. This paper seeks to kick-start more of these conversations.

Finally, we should recognise how much we have already achieved in applying data science to problems in healthcare. As a senior Māori health advisor put it, we should "build a culture of sharing our successes and really talking about them. There are not enough grassroots champions of data science. Data science is now, not a future thing."