

Emerging Health Technology Introductory Guidance

January 2019
V0.4



**For safely developing & using
Algorithms in Healthcare**



Purpose and background

Purpose

The Introductory Guidelines for safely developing and using Algorithms in Healthcare document intends to provide good practice advice on the governance, development and application of algorithms for individuals and organisations working in the New Zealand Health System. It provides information on the principles and recommendations that should be considered to ensure the best outcomes while also aligning to the Privacy Act and Health Information Privacy Code.

Scope

This document focusses on four key themes, their challenges and good practice guidelines for the creation, introduction and management of new algorithms in the New Zealand health sector. It is not yet intended to be retrospectively applied for existing algorithms, however this should be considered in the future.

Background

The Ministry of Health held a workshop in late August 2018 to discuss the safer and more effective use of algorithms in the health sector. The workshop included attendees representing government agencies, public sector, health professionals as well as industry and regulatory experts.

Presentations were given by attendees on ethics, research, industry and sector view, followed by a round-table strategy session with participants focusing on the four topics of Governance, Bias, Operationalising and Assurance when developing algorithms for the health sector.

This document summarises the key points raised during these discussions, adding additional detail and use cases to highlight specific issues. The workshop focused on a future view as the prevalence of algorithms to support health care decisions increases within the New Zealand health sector and worldwide.

The intended audience is people or organisations planning to develop implement or use algorithms in the planning and delivery of healthcare in New Zealand.

Why is a framework important? Why now?

The amount of data available within the New Zealand Health System is growing quickly, as is the ability to capture and record massive amounts of information about an individual patient and their health journey.

In parallel, we are seeing advancements in technology, computational capacity, and access to sensors which generate copious amounts of machine readable data. This means our ability to generate and develop novel and improved algorithms has gone from a very limited and specialised field, to one which is becoming embedded in every day tools and tasks.

Despite this, algorithms are not new to health care. They are being used on previously unused data to develop useful insights, which in turn are helping improve operational efficiency and support clinical decision tools to provide better and safer care. However the changing technology environment means how algorithms are being developed is changing, especially as we move into an age of Artificial Intelligence and Machine Learning.

This situation presents great opportunities for the New Zealand healthcare sector, but as this often relies on data that the Ministry of Health, DHBs and health care providers have strong ethical duties with regard to data. In this we mean that data-driven technologies must be harnessed in a safe, evidenced and transparent way.

It is important to preserve appropriate human oversight and ensure that the views of key stakeholders, notably the people who will receive or participate in services, are given the appropriate consideration.

Finally our approach to honest and open communication means that if there is a risk that an algorithm could have unintended consequences for patient or population harm, we must ensure that the creation and deployment is rigorously examined and tested.

This is commonly known as social licence. While this document doesn't focus on this as specifically, is an overarching topic that will be addressed more comprehensively in the creation and implementation the Ministry's Data Strategy.



Purpose and background

What are algorithms?

Algorithms are the automatic decision-making processes used by computer programs to identify patterns in data. They have an essential role in supporting the services government provides, and help deliver new, innovate, and well-targeted policies for New Zealanders. (data.govt.nz)

McKinsey gives a good breakdown of the differing aspects:

Machine Learning – the ability of computers to detect patterns in large data sets through the application of algorithms.

Artificial Intelligence – the science and engineering of automated problem solving. The object is to generate solutions by using computers to mimic the cognitive functions associated with deliberative thought, including perception, reasoning, and learning.

Predictive Modelling – a machine-learning approach that builds pattern-recognition models using sample data with known attributes and outcomes (labelled ‘training data’). Working from known patterns, the model can predict outcomes for new observations. Machine-learning magnifies the power of predictive models through great computational force.

Deep Learning – the most advanced technique for predictive modelling. Connects software-based calculators to form a complex artificial ‘neural network’ often many layers deep.

Natural Language Processing – a sub-field of artificial intelligence that is focused on enabling computers to understand and process human languages, to get computers closer to a human-level understanding of written and spoken language.

Algorithms are currently used throughout the New Zealand health system with examples such as:

- assisting clinical prioritisation (CPAC)
- patient assessment (InterRAI) outcomes
- [University research](#) into natural language processing software to identify GP consultations for Zoster

What’s happening in New Zealand?

In May 2018 the Government Chief Data Steward and the Privacy Commissioner jointly published six principles for the [safe and effective use of data and analytics](#) by government agencies.

- Deliver clear public benefit
- Maintain transparency
- Understand the limitations
- Retain human oversight
- Ensure data is fit for purpose
- Focus on people

This was followed by the publishing of the [Algorithm Assessment Report](#) in October 2018 which presented an assessment of information reported by 14 government agencies about the computer algorithms they are using to deliver functions. It focused on areas where algorithms are used in decision-making processes that affect people in direct and significant ways.

Recommendations from this report targeted the following areas:

- Human oversight
- Development and procurement
- Information and transparency
- Review and safeguard
- Sharing best practice

IoT Blood Pressure Variability Study

Healthcare is using newly accessible data to discover new insights about traditional assumptions such as the variability of blood pressure in a population. This was historically assumed to centre around 120/80mmHg.

Key Points

- 17 million measurements
- 56,000 participants
- 6 months
- no extra cost

Provides reference values for future understanding of the nature of BPV and help guide individualised management

http://static.withings.com/content/whi/nl/dec16/poster_bpv.pdf



Purpose and background

Our Areas of Focus

As a part of engaging with the sector, and the development of the StatsNZ led work, the Ministry of Health identified that the health and disability sector would benefit from guidance to support those undertaking algorithm development and its subsequent use in the New Zealand health sector. In this context, we identified four concepts for discussion at the August 2018 workshop and the guidance aligns to these. The concepts discussed were:

Governance

- How do you ensure that algorithms are well governed?
- Are there reporting tools to understand all the previous elements?

Managing Bias

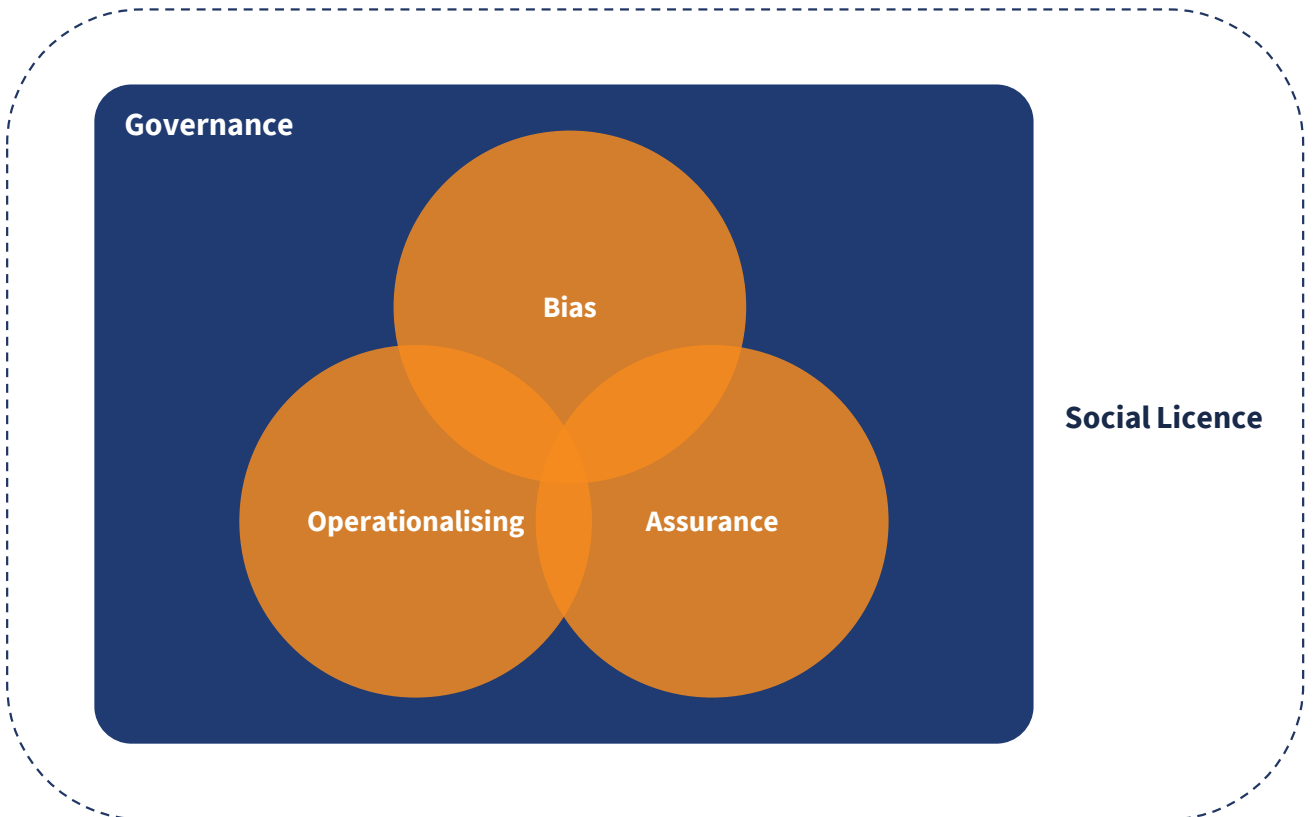
- Why is this needed?
- Tools or processes to be used in assessing bias.

Operationalising Algorithms

- What roadblocks need to be considered?
- Principles or approaches for addressing these.

Assurance/Confidence

- How do you know that this is a good algorithm – how do you have confidence?



Governance

New Zealand has many unique attributes within its population and healthcare setting, which need to be governed accordingly. We are an ethnically diverse nation, with a publicly funded healthcare system. The NHI (National Health Index) allows a longitudinal view of a patient's journey through this system, which in turn allows the Ministry of Health to create robust national datasets.

When an algorithm's outcome can make or support decisions that affect a person's life, a natural regulatory response is to demand oversight and hold people accountable if something goes wrong based on the output. The importance of a strong strategy and a diverse governance structure will help mitigate this.

Healthcare organisations are used to working in highly regulated technology environments (e.g. pharmaceuticals and devices). However software, unlike most areas of healthcare, is an minimally to non-regulated technology. Because of this, the organisational and population risk/benefit outcomes are not nearly as well defined as other capabilities. For example, adoptors or developers of AI in healthcare settings need to be aware that unlike drugs or devices, these tools have not been rigorously evaluated against an international standard.

SaMD (Software as a Medical Device) embedded in devices undergoes some regulatory hurdles, but software for things such as hospital logistics and decision tools are currently unregulated within New Zealand.

An effective governance structure should build confidence in, and demand for, new data-driven innovations amongst private and public sector and New Zealanders.

Governance is a shared opportunity and a shared responsibility for all parties involved in the creation, use, assessment and reviewing of algorithms (and its referenced data) used in a health care setting. The purpose of governance is to ensure strategic alignment and preserve the effectiveness, and reduce undesired outcomes from their implementation.

This overarching structure allows mitigation of the challenges and promotion of good practice when focusing on the key areas of Bias, Operationalising and Assurance.

Where do the challenges lie?

No algorithm is perfect, but the question is whether it is better than the next option. For example – is automating a process with the potential of a network outage or scheduled downtime maintenance, better than the current process which relies on a fax machine not 'running out of paper' therefore risking catastrophic data loss?

In order to understand this, understanding the context and impact of the algorithm is key. There is a real difference between an algorithm for fraud detection vs one that decides whether someone gets treatment or not (lives or dies). Understanding the risk level of implementation, and understanding the limitations of the algorithm itself as a decision making tool is critical to effective to good governance of algorithms.

The difficulty for governance faced with this situation is that there is often a disconnect between the source of the data, those developing the algorithm and those who will be impacted.

This often leads to governance delaying progress to understand all the elements of the algorithm before it can proceed to the next phase.

A governance approach also needs to consider the issues of social license they are utilising to progress the development of an algorithm. This includes thinking about the privacy implications of the source data and algorithm, the legal and ethical framework around the source data used to develop the algorithm, the security and any anonymity of the data used.

Finally the governance approach needs to be conscious of using algorithms to more equitably make decisions or allocate resources. A strategic question that needs to be answered is whether an algorithm that references historical data may have inherent bias (eg. previous under selection) which in turn has the impact of further perpetuating the current inequitable situation.

What does good practice look like?

The first focus is the people involved in overseeing the governance, ensuring that the right people are involved.

Governance

Risk can be mitigated by including stakeholders from across a diverse and multi-disciplinary setting, with a variety of skills and multi-dimensional input and view. This includes those who collect the data, those who build algorithms and those who are impacted by algorithms or those who are accountable for the service changed by the algorithm.

It is also important for this group to learn continuously and to raise the general knowledge of all members. Good governors learn as much as they can about what they're governing so that they can make better decisions and can make effective suggestions to the people being governed. It is encouraged that we share learnings, experiences, process between parties and agencies.

The principles of partnership, participation and protection underpin the relationship between the Government and Māori under the Treaty of Waitangi. It is important to consider how the development of an algorithm conforms with the expectations from the Treaty of Waitangi and expectations from Māori regarding Data Sovereignty, and the opportunity that an algorithm can present for improving equity and outcomes for Māori.

Focus on decisions to implement algorithms that will deliver clear public benefit over other options – whether at a population level or to address inequalities for a specific group, it is a good way to build confidence in the process of developing algorithms.

This along with a clear and transparent governance framework and decision making process will build confidence that the intent is just. The framework should include privacy and security regulation and Privacy Impact Assessments which should be completed at an early stage.

Further to this, especially if there may be concerns where there might be marginal impact, an independent ethical review could be used to provide some reassurance to the governance approach.

Key Points:

- Make sure the right people are involved in overseeing governance
- Include stakeholders from across a diverse and multi-disciplinary setting
- Have a clear and transparent governance framework
- Focus on decisions to implement algorithms that will deliver clear public benefit over other options
- Understand the context and impact of the algorithm
- Understand the risk of implementation
- Understand the limitations as a decision-making tool
- If the algorithm references historical data, does it have inherent bias?
- Think of the privacy, legal and ethical implications around the source data used to develop the algorithm
- Ensure the algorithms conforms with the expectations from The Treaty of Waitangi and expectations from Māori regarding Data Sovereignty
- Conduct Privacy Impact Assessments at an early stage
- If necessary, commission an independent ethical review
- Learn continuously and raise the general knowledge of all members
- Share learnings, experiences and process between all parties and agencies

Areas of Expertise within Governance Structure

- Methodological – understanding the methodology, both within governance and data science
- Data Structure – understanding the data itself – collection, interpretation and use
- Organisational Strategy – the environment in which the algorithm will be implemented
- Clinical – the health professionals using and interpreting the outputs of the algorithms
- Advocacy – representatives of the groups the algorithm will be used on

Case Study – New Zealand PREDICT Study

New Zealand PREDICT Study

Cardiovascular risk assessment in New Zealand has, until 2018, been based on the Framingham cardiovascular risk charts. These were developed in the 1960s and 1970s from the Framingham cohort study in the United States, and allow clinicians to calculate a patient's future risk of cardiovascular disease by taking into account factors such as blood pressure, cholesterol levels and smoking status. These equations still provide a reasonable approximation of a patient's risk, their limitation, however, is they do not take into account New Zealand's ethnic diversity, and may under- or overestimate risk in some patients. In addition, since the time of the Framingham study, more cardiovascular risk factors have been identified, such as fasting blood glucose levels or HbA_{1c} and renal function.

The PREDICT study is a New Zealand research project which began in 2003 with the aim of deriving cardiovascular risk prediction equations based on local data. By December 2015, approximately 400,000 patients aged 30–74 years had been assessed. The results of the PREDICT study have been used to develop the NZ Primary Prevention equations, which now form the basis of CVD risk assessment in New Zealand. These equations incorporate more variables than the Framingham equations, in order to improve the accuracy of prediction and therefore help clinicians to provide appropriate targeted care.

The NZ Primary Prevention equations are becoming available for clinicians to use in practice, and the recommendations in the 2018 CVD risk assessment consensus statement based on these equations can be applied now.

PREDICT is a web-based decision support system, used mainly to assist primary care practitioners to assess and manage cardiovascular disease risk. It has been developed by a research team at The University of Auckland, and software company Enigma Publishing Limited.

Periodically, the University of Auckland (through the 'VIEW' research group) obtain an anonymised PREDICT data extract from those organisations who have opted in to collaborate with their research. This includes data from 55,000 Māori, 55,000 Pacific people and 35,000 people of Indian descent. In addition, data has been made available through Statistics NZ's Integrated Data Infrastructure (IDI). This is administration and government data that has been linked to anonymised individuals' health records, which have been made available for public good research.

In 2015 the Ministry of Health commissioned the Heart Foundation to review the relevant evidence on CVD risk assessment and management. As part of this review, the University of Auckland VIEW research group provided the Heart Foundation with two pre-publication papers describing new CVD risk prediction equations. Seven areas identified for review were:

- the expected real-world benefit to New Zealanders of having New Zealand-specific risk stratification and risk equations
- the CVD risk assessment window or frequency for different risk categories
- the evidence for medication treatment thresholds and goals of treatment
- lifestyle interventions, including dietary advice that is sustainable for populations with health literacy challenges
- effective ways to encourage those at increased CVD risk to change their behaviour in a sustained way and take their medication, including through effective risk communication, shared decision-making and goal setting
- co-morbidity with serious mental illness, the increased risk linked with serious mental illness, and impact of antipsychotic medications
- overall consistency of New Zealand guidelines with new international guidelines.

Bias

Algorithmic bias occurs when a computer system reflects the implicit values of the humans who are involved in the coding, collecting, selecting or using data to train the algorithm.

Recognition of how previous health services were delivered, and the bias inherent in this means the use of historical data does not necessarily reflect the of New Zealand healthcare outputs going forward.

Bias can be introduced to an algorithm in several ways. During the development and implementation of an electronic data collection, data must be collected, digitised, adapted, and entered according to human-designed cataloguing criteria. Programmers then assign priorities for how a program assesses and sorts that data. This requires human decisions about how the data is categorised, and which data is included or discarded.

Though well-designed algorithms frequently determine outcomes that seem equally (or more) equitable than the decisions for human beings, bias still regularly occurs. But if algorithmic systems are at least partial products of human judgments, assumptions, simplifications and curatorship, can they ever be truly neutral and fair?

Types of Bias

Type	Description
Pre-existing	A consequence of underlying social and institutional ideologies. Such ideas may influence or create personal biases within individual designers or programmers. These can be explicit and conscious, or implicit and unconscious.
Technical	From limitations of a program, computation power, its design, or other constrain on the system.
Emergent	Algorithms that may not have been adjusted to consider new forms of knowledge (e.g. new drugs or medical breakthroughs, new business models).
Correlations	Unpredictable correlations can emerge when large data sets are compared to each other.
Unanticipated uses	Emergent bias can occur when an algorithm is used by unanticipated audiences. Reliance on the software instead of in partnership with their own knowledge can indirectly lead to bias by narrowing potential pathways.
Feedback loops	Emergent bias may also create a feedback loop, or recursion, if data collected for the algorithm results in real-world responses which are fed back into the algorithm.
Intentional	Deliberately prioritising data in order to influence results in a predetermined direction.

Where do the challenges lie?

The concept of a “half-life” for clinical data implies that more recent data is better than more data when predicting the future. This suggests that prioritising smaller amounts of recent data may be more effective than using larger amounts of older data towards future clinical predictions. If this is an acceptable approach to mitigate bias, could older data be used if weighted for it’s likely bias?

The difficulty in attempting to mitigate bias is in part the difficulty in observing it, similar to the difficulty encountered when attempting to “prove a negative”.

Combined with the fact that an algorithm in use will begin to affect the source data that was used to create the algorithm means that the future will look like the outcome selected for in the algorithm.

As is considered with the peer review and publishing of scientific journal articles, it is likely that the motivation of the algorithm developer is influenced by the interests that the developer has – essentially designing outcomes that suit and then finding inputs to create the outcomes.



Bias

What does good practice look like?

Acknowledge that bias exists, and take the appropriate steps to minimise its impact on the algorithm outputs. When working with factors such as age, gender, or ethnicity it is important to incorporate them whilst also addressing the social bias that may occur from these particular attributes within the algorithm code itself. Ensure equity to the access and outcomes produced by the algorithm's results. [\[LINK\]](#)

Consider in the algorithm that there may be diminishing returns for including additional data elements. If there is an ability to reduce the number of elements, for no appreciable difference in performance, this may increase how understandable the algorithm is. Also consider automated fields (eg. date/time stamps, the completion rate, and the variability of collection) for use in your algorithm.

Ensuring that the uncertainties that arise in the algorithm output are duly accounted for in the decision-making and governance process and appropriate steps have been taken to mitigate unwanted side effects means that users can be assured of which bias has been identified and addressed.

Understand there are two groups of patient data – patients whose data is used to develop the algorithm and patients on whom the algorithm data is used. The approach should be to develop and deploy an algorithm based on the New Zealand population when possible, rather than adopting an algorithm designed by another jurisdiction.

Being able to explain the algorithm, the impact, the logic and process, and provenance is critical. Have an auditable design and methodology to ensure the review process includes a diverse and multidisciplinary collaboration for best visibility.

Follow applicable data standards (eg. SNOMED) or conventions – per applicable data dictionaries (eg. National Minimum Dataset) to ensure that the data is being used for the purpose it was collected (this may help ensure the outcomes deliver those which you might expect).

Key Points:

- Acknowledge bias exists! E.g.
 - Method of previous healthcare delivery
 - Historical data collection – including digitisation and categorisation
 - Data prioritisation
 - Human judgements, assumptions and simplifications
 - ‘Half-life’ of clinical data & weighting
 - Data output from algorithm in turn becomes source data
 - Personal motivation in development
- Take appropriate steps to minimise bias impacts on the algorithm outputs
- Difficulty in attempting to mitigate bias is in part the ability to observe it in the first place
- When working with factors such as age, gender, ethnicity it's important to incorporate them whilst also addressing the social bias that may occur
- Are the additional data elements necessary? Reducing the number of elements may increase the algorithms understandability
- Ensure uncertainties in the output are duly accounted for, and steps have been made to mitigate unwanted side effects
- There are two groups of patient data – patients whose data is used to develop the algorithm, and patients on whom the algorithm data is used
- Be able to explain the algorithm and its impact. Be auditable.
- Follow applicable data standards

Case Study – Allegheny Family Screening Tool

Excerpts from [Developing Predictive Risk Models to Support Child Maltreatment Hotline Screening Decisions](#) - Allegheny County, Pennsylvania, USA

Predictive Risk Modelling (PRM) uses routinely collected administrative data to model future adverse outcomes that might be prevented through a more strategic delivery of services. In the context of child protective services, PRM tools can be used to help child protection staff make better initial screening and service decisions for children who have been named in reports of alleged abuse or neglect.

In August 2016, the Allegheny County Department of Human Services (DHS) implemented the *Allegheny Family Screening Tool* (AFST), a predictive risk modeling tool designed to improve child welfare call screening decisions. The AFST was the result of a two-year process of exploration about how existing data could be used more effectively to improve decision-making at the time of a child welfare referral.

The process began in 2014 with a Request for proposals and selection of a team from Auckland University of Technology led by Rhema Vaithianathan and including Emily Putnam-Hornstein from University of Southern California, Irene de Haan from the University of Auckland, Marianne Bitler from University of California – Irvine and Tim Maloney and Nan Jiang from Auckland University of Technology. Input was solicited throughout the exploration and development process and used to inform the final product. Prior to implementation, the model was subjected to an ethical review by Tim Dare of the University of Auckland and Eileen Gambrill of the University of California-Berkeley.

In mid-2015, it was decided that the most promising, ethical, and readily implemented use of PRM within the Allegheny County child protection context was one in which a model would be deployed at the time an allegation of maltreatment was received at the hotline. The objective was to develop a decision aid to support hotline screeners in determining whether a maltreatment referral is of sufficient concern to warrant an in-person investigation.

It should be noted that while in some settings machines have been used to *replace* decisions that were previously made by humans, this is not the case for the Allegheny Family Screening Tool (AFST). It was never intended or suggested that the algorithm would replace human decision-making. Rather, that the model should help to inform, train and improve the decisions made by the child protection staff.

The approach that Allegheny and the research team have taken to the implementation of the Family Screening Score is to see it as a three way evolution between practice, policy and modelling. Because practice and policy is evolving, the best way to build and implement the model will also change. At some point, they would expect this process to settle into a more stable equilibrium.

Particular issues considered during the Ethical Analysis were:

- Consent
- Information about other family members
- False Positives/False Negatives
- Stigmatization
- Racial Disparity
- Professional Competence/Training
- Provision and identification of effective interventions
- Ongoing monitoring
- Resource Allocation

“As we emphasized throughout the Ethical Analysis, decisions are being made right now. It is not a matter of making or not making related decisions. The decisions involved are complex ones made in a context of inevitable uncertainty that contributes to inevitable error. Research on decision-making in the helping professions highlights the play of biases and fallacies. Confirmation biases are common in which we seek information that corresponds to our preferred view (e.g., there is no abuse) and fail to seek evidence that contradicts preferred views. Errors of omission (failing to act) are viewed as less harmful than errors of commission (acting - for example, removing a child from the care of her family). The question is, how can we make the fewest errors in our efforts to protect children and families? AFST seems an ethical and potentially important contribution to that effort.”



Operationalising

From research to roll-out - there is a huge jump from doing the research/pilot to using the algorithm to make real decisions in a clinical practice or an operational setting.

Where do the challenges lie?

Not having access to good data is a critical issue. Researchers are often given data sets from a static period of time, which has been well maintained, is well understood and is of generally high quality. This means that when the algorithm is run over a production set, results can be very different to the current operational situation. Research data in this context needs to be relevant and available, considering what you will know at the time of the algorithm generating outputs (i.e. some data which is used as predictor is generated after the event).

There is also often a disconnect between the technical development and the clinical process. This is noticeable in larger projects which may involve multi-site collaboration. What needs to be considered and understood in the project is what the real world problem is that is being supported, including how it will be used.

In a clinical setting a binary “yes or no” can be confronting to groups who regularly synthesise risk using “[clinical judgement](#)”, an often poorly documented and intuitive process. This means most practitioners find a relative risk or risk score more

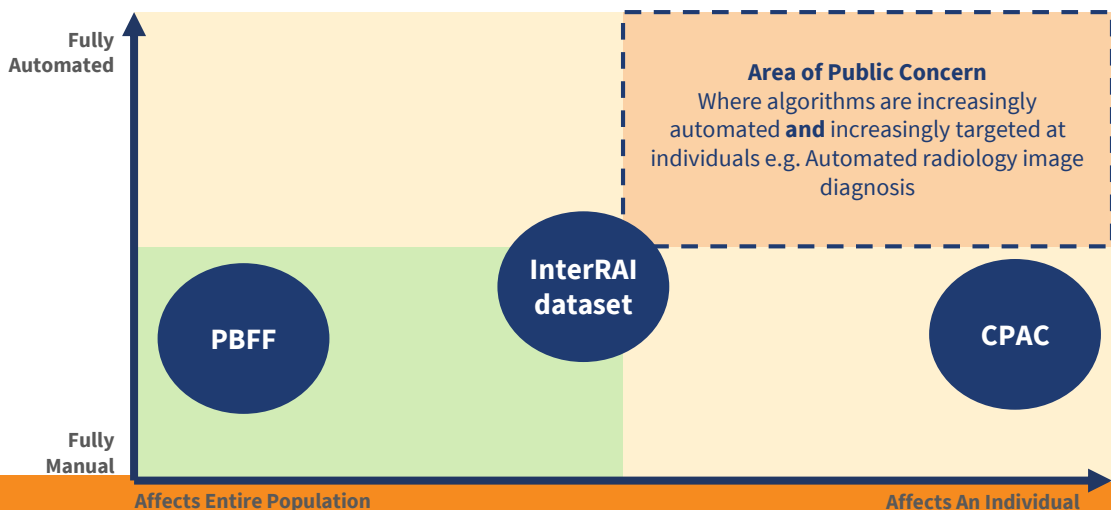
acceptable as they then use this as an input to help with making a clinical judgement. While it is important to integrate an algorithm with current tools, there is a need to consider that this may mean inheriting bias from existing processes and databases.

Commonly these projects will begin with funding for the research, followed by more funding being sought for operationalising the algorithm. However at this time funding must also support the algorithm through its full lifecycle to ensure improvements and reviews to functionality remain available through the useful life of the algorithm.

The final element which is a large challenge is the engagement with those who are affected by the algorithm. There have been a number of examples of where seemingly useful algorithms developed using a solid methodology have, once they have moved from research to implementation, been abandoned as the developers have taken insufficient account of the impact the algorithm will have on the subjects.

[Social license](#) describes an organisation’s or project’s legitimacy, credibility and trust in the eyes of the public or key stakeholders. Below is a conceptual framework for thinking about how algorithms trigger concerns with social license and the need to engage. It should be expected that engagement increases significantly as the algorithm moves towards the top right quadrant. The risk of not engaging will increase the likelihood of a public backlash when attempting to implement a new or improved algorithm.

Conceptualising Social License for Algorithms



Operationalising

What does good practice look like?

Begin with understanding the use case – ask about the reason for the optimisation, and ask whether an algorithm is the right answer. This will ensure that the algorithm is focused on a real business problem.

Working directly with stakeholders to understand the problem will ensure that this is addressed up front. This needs to be coupled with an understanding of the potential impact of the algorithm. It is highly likely that a large and positive impact on the problem is one which the organisation and staff will find useful and support, and will help garner social license with those who are affected.

Connecting the business problem with the data will ensure that the right datasets have been made available and understanding the business process and the use case are represented in the data. This approach is a way of describing who should be involved, having the right mix of business, technical and policy expertise within the development team is the right starting point.

The algorithm should always be tested in a clinical or operational environment – whether as part of clinical trials, or silently in a testing capacity on production data. The closer the test data is to real data, the more likely it is to be effective e.g. including non-completion data (null fields) is just as useful as other data.

Consider simulation testing with the users, see how it fits with the clinical or business process and how it could be deployed. A further consideration is the link between training and test data. There is a need to maintain separation between these, with data for testing performance only used once, to avoid over-fitting or accusations of generating unrealistic results.

Create and maintain a robust set of technical documentation so that someone can reproduce what you have done is important, particularly in the context of building social license and for assurance and governance processes. We need people to be able to reproduce our algorithms as one form of quality control.

There is also a need to build commitment to training and data literacy for the users and business owners. This buy-in and understanding for the end-users, including any staff who will be using an algorithm's outcome is important to build social license. There will be a cultural shift where users may go from using the data they see, to using the algorithm outputs, and these users need to understand the risks and benefits. This will also provide a broader engagement point for the public or people who are affected by the algorithm's outputs.

Next, but probably most important is the communications and engagement approach. The project should prioritise how the operationalisation of the algorithm will be documented, communicated and discussed with those who will be affected. This includes community groups, advocacy organisations and within New Zealand, a partnership, participation and protection framework to engagement with Māori. This will include how the solutions will improve equity and how it maintains tino rangatiratanga.

In order to better engage the public, some processes and tools which are of use are:

- [Ethics approval](#) - to assure people that the benefits outweigh the risks
- Privacy impact assessment - to identify the impact the project might have on the privacy of individuals involved
- Public communications or community meetings - depending on the social license required there may need to be more.

The final element to consider is conducting the work in the open, an approach not common in healthcare (often due to privacy concerns) however it can be very reassuring to the public if the group is transparent in its approach and implementation.

Lastly there is a need to share processes, experiences and lessons learned between all of those engaged in algorithm development. Given the rapid increase in availability of more powerful and easy to use tools, we are all required to hold each other to account to ensure the collective social license accrued for this use isn't eroded through rogue practice.

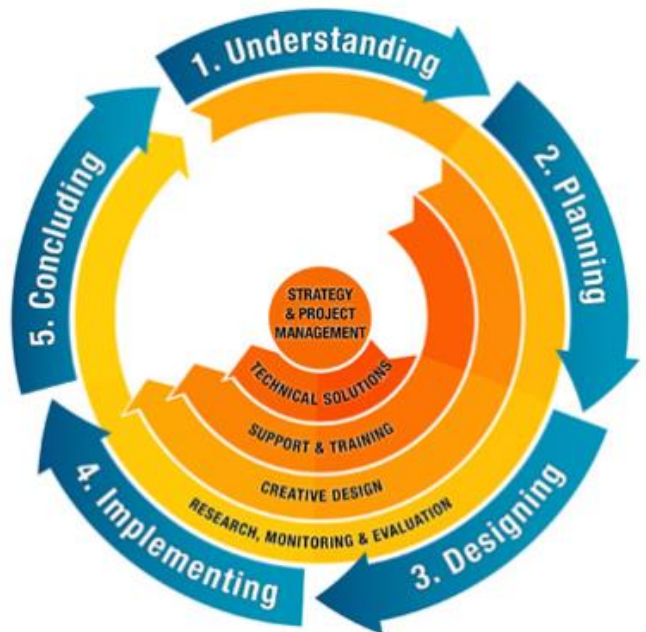


Operationalising

Key Points:

- Understand the use case – the reason for the optimisation, and whether an algorithm is the right answer
- Connect the business problem with the data to ensure the right datasets have been made available
- Understand the business process
- Engage early with those who will be affected by the algorithm's implementation to understand its impact
- Have the right mix of business, technical and policy expertise within the development team
- Datasets used in research and development need to be relevant and available, and as close to production sets as possible
- Test the algorithm in a clinical or operational environment – whether as part of clinical trials or silently in a testing capacity on production data
- Non-completion data is just as useful as other data
- Create and maintain technical documentation so that someone can reproduce what you have done
- Ensure project funding covers not only the research and operationalising, but the full lifecycle (improvements, maintenance and review) of the algorithm
- Commit to training and data literacy for the users and business owners
- Have a communications and engagement plan, ensure all parties affected have been involved and kept updated

- Does this algorithm need ethics approval?
- Complete a privacy impact assessment
- Consider an open and transparent approach to development and implementation
- Share processes, experiences and lessons learned between all of those engaged in the algorithm development



Case Study – Virtual Diabetes Register (VDR)

**This use case highlights the importance of revisiting an algorithm's outputs, retesting and modifying against latest research and operational results. An algorithm's lifecycle does not end when it goes live in a production/clinical environment.*

Each year the New Zealand Ministry of Health releases national estimates of the prevalence of diabetes based on the **Virtual Diabetes Register (VDR)**. The VDR is an important tool to monitor prevalence of diabetes and support national and local clinical quality improvements.

It contains data about people suspected as having diabetes, identified through their use of diabetes health services. The VDR uses an algorithm to identify these people in data extracted from inpatient, outpatient, laboratory test and pharmaceutical dispensing data collections. The Register is collated annually at the end of March and national and regional diabetes prevalence estimates are calculated based on the number of people on the VDR as at 31 December of the previous year. People with diabetes who were deceased and those not enrolled in a PHO were excluded from the totals.

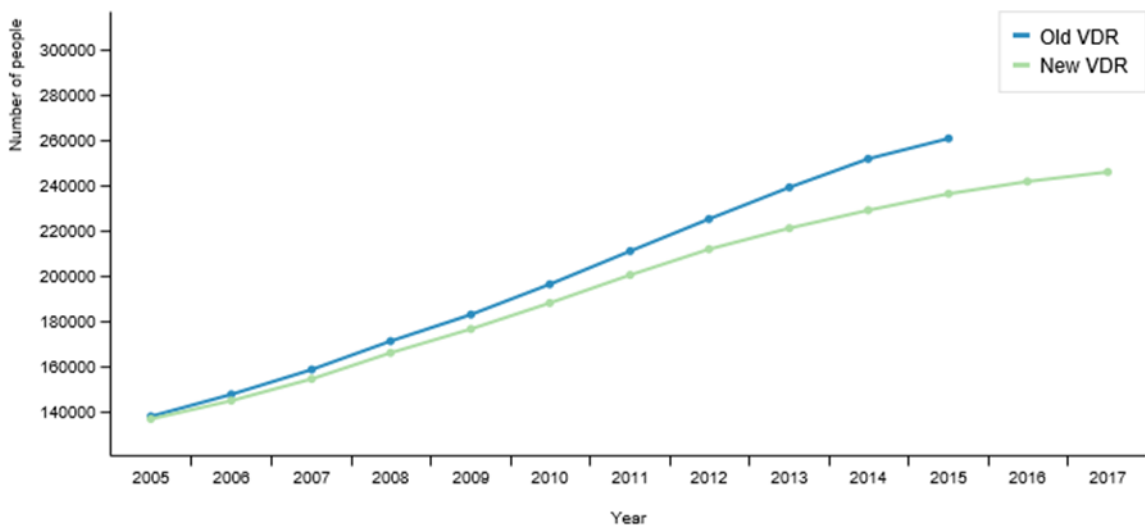
2017 revision

In 2016 the algorithm used to create the VDR was assessed against the Auckland TestSafe repository of actual glycaemic test results. (Paper - [Can administrative health utilisation data provide an accurate diabetes prevalence estimate for a geographical region?](#)).

The diabetes prevalence estimate based on the original 2014 MoH VDR was 17% higher than the corresponding TestSafe prevalence estimate. Compared to the diabetes prevalence based on TestSafe, the original VDR has a sensitivity of 89%, specificity of 96%, positive predictive value of 76% and negative predictive value of 98%. The modified VDR algorithm has improved the positive predictive value by 6.1% and the specificity by 1.4% with modest reductions in sensitivity of 2.2% and negative predictive value of 0.3%. At an aggregated level the overall diabetes prevalence estimated by the modified VDR is 5.7% higher than the corresponding estimate based on TestSafe.

As a result, improvements to the algorithm were made in early 2017 to create the latest version of the VDR. The comparison highlights the potential value of a national population long term condition register constructed from both laboratory results and administrative data.

Number of people estimated as having diabetes, on the old VDR and the new VDR, 2005–2017



Assurance

People want to know that their privacy and rights are safeguarded and to understand how and when data about them is shared, so that they can feel reassured that their data is being used for public good, fairly and equitably.

Before using an algorithm, the project needs to identify the number of people it will affect, how and when the data informing the algorithm will be collected, whether the algorithm will make recommendations or decisions, and whether the algorithm's output can be audited.

During development and once the algorithm is in use there is needs to be documented assurance that the algorithm is conforming to the outputs that were planned.

Where do the challenges lie?

Ensuring the algorithm applies in the context it was designed for is the key challenge. If it was designed for X but being used for Y, then the output from the algorithm could put people at risk. There needs to be confidence that an algorithm does what you think it does.

Understand where the algorithm will take you once it is implemented e.g. moving from manual to automated process using machine learning – or from population level focus to that of an individual. There needs to be an understanding of this change in focus and assurance that the movement is consistent with the outcomes you are trying to achieve.

A methodology should be used, but the key question to be asked is *what does success look like?* It is necessary to define the levels of accuracy of the algorithm – especially in the context of personal health vs public health outcomes. The measure of accuracy and therefore success will depend on whether the algorithm output delivers the intended outcomes within the tolerances required.

Another key element of the methodology to be considered is the standard of data that is available, and if the data quality changes, when is it deemed to be insufficient for achieving the outcomes required.

The risks and benefits of scalability need to be identified and understood if there is intent for the

algorithm's scope to be expanded. Algorithms in clinical settings can be used to augment or even replace human cognition. However, because they can scale so easily, in addition to the benefits being magnified, the risk of a bad algorithm causing the equivalent of a plane crash should also be a real concern.

Lastly there is an obligation to think about how you manage any interests as a part of the algorithm development and delivery, with a keen eye on the potential for conflicts.

What does good practice look like?

Making a strategy for risk management can involve more than just deciding whether to accept risk or not. If the algorithm is part of a bigger business process, understanding the risks and benefits of implementation can help spread the risk across a number of areas. By spending time and resources on a risk management strategy, it can reduce the chances of negative impacts and potential harm. A higher risk project may require more pilot time and more assessment than one with a lower risk.

Retain human oversight, ensure that the governance approach regularly receives reports about the algorithms ongoing effectiveness at achieving the intended outcomes, while also informing any smoke signals you might expect if the algorithm is delivering unacceptable outputs.

Immutability and a history of the results is very useful to identify when an algorithm changes course. Using a standard test set which is routinely run across the algorithm will ensure that expected outputs are consistent with actual outputs.

A common approach to provide assurance in healthcare is the use of peer-review. In this case using other organisations in the health sector will ensure that the team doesn't suffer from tunnel vision.

Developing a regular schedule of this oversight, testing and peer review is best practice. Maintaining transparency and accountability with clear operational principles to explain decisions and outcomes following the implementation is the best approach. This is often done in accounting with independent auditors.



Assurance

Key Points:

- Identify the people who will be affected by the algorithm's use
- Understand how and the data informing the algorithm will be collected
- Ensure the algorithm applies in the context it was designed for
- Adopt the appropriate algorithmic methodology
- Understand the benefits and risks of scalability
- Manage any interests as part of the development and delivery, and identify any potential conflicts
- Make a strategy for risk management at an early stage
- Retain human oversight
- Keep governance updated with progress and identify and escalate any risks
- Maintain a regular schedule of testing and peer-review
- Maintain transparency and accountability with clear operational principles to explain decisions and outcomes



Case Study – Population Based Funding Formula (PBFF)

**This use case highlights the importance of revisiting an algorithm’s outputs, retesting and modifying against latest research and operational results. An algorithm’s lifecycle does not end when it goes live in a production/clinical environment.*

The **Population-Based Funding Formula (PBFF)** is a technical tool used to help equitably distribute the bulk of the district health board funding according to the needs of each DHB’s population.

The formula takes into account the number of people who live in each DB catchment, their age, socio-economic status, ethnicity, and sex. It also has mechanisms to compensate DHBs who service rural communities and areas of high deprivation.

The funding covers a range of health services including primary care, hospital and community care, health of older people, and mental health.

The aim of the PBFF is to equitably distribute available funding between DHBs according to the relative needs of their populations and the cost of providing health and disability support services to meet those needs. The PBFF gives each DHB the same opportunity, in terms of resources, to respond to the needs of its population.

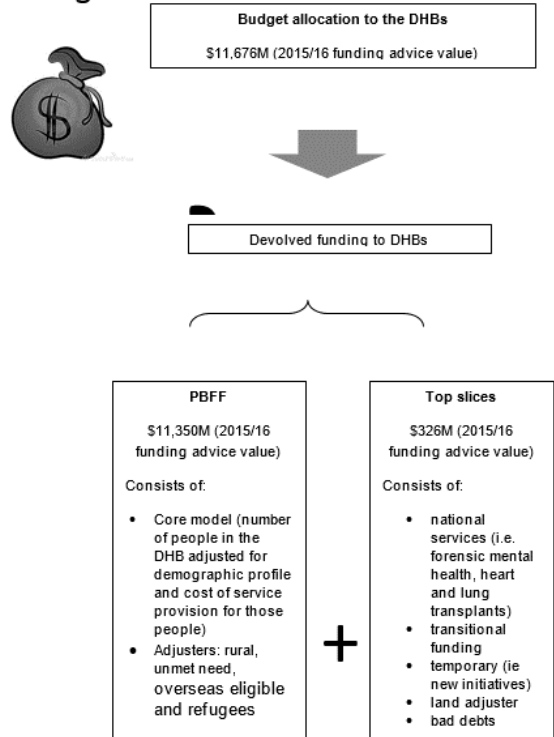
The PBFF provides a number of advantages for both DHBs and the Government, which include that it:

- allows for an equitable allocation of DHB funding based principally on local population needs
- maintains per head level of service
- allows for greater responsiveness to changing population needs
- allows for specific adjustments to reflect some of the unique costs of providing health services across the country
- it is a technical model that avoids protracted individual bidding and negotiation and lobbying
- promotes efficiency and fiscal control for the government, and
- places responsibility for managing within the Vote on the Minister of Health and frees up Cabinet to consider the major strategic issues facing Health.

For the 2014/15 review a Technical Advisory Group (TAG) was set up with members from DHBs, the Ministry of Health and Treasury. Overall, the changes in the formula were minimal and the consistency between results with respect to the new cost weights highlighted the robustness of the model over time.

The PBFF is currently reviewed on a five-yearly cycle.

Budget overview



Ethics & Algorithms Toolkit Checklist

This is checklist released in the US to help cover the ethical and governance questions that should be considered in the development of algorithms. Below are a selection of questions that relate well to the health sector
Sourced and adapted from: www.ethicstoolkit.ai

Overview	Description
Who is the toolkit for?	For anyone building or acquiring algorithms in the government sector and beyond
What is the toolkit?	A process to walk the user through a series of questions to help understand the ethical risks posed by using an algorithm, and identify what you can do to minimise those ethical risks.
Who made this toolkit?	The beta release was a collaboration between The Center for Government Excellence (GovEx) at Johns Hopkins, the City and County of San Francisco, Harvard DataSmart, and Data Community DC.

1. People

Have you identified all the stakeholders? (the person affected, the people that use)

2. The data

Do you fully understand the dataset and how the data was collected?

Does the context of the collection match the context of your use?

Who is represented in the data? Who is under-represented or absent?

Would the use of the data surprise subjects?

Are there any fields that should be eliminated from your data?

3. The risk level

Can you describe to what extent this algorithm would impact on an individual or the population?

4. The methodology

Can you describe the logic that connects the variable to the output of your equation?

How did you determine what weight to give each variable?

What assumptions are you relying on to determine the relevant variables and their weights?

Have you determined whether your outcome expectations makes sense to a diverse audience?

5. The bias

Do you have someone in your team tasked specifically with identifying and resolving bias and discrimination issues?

Will your variables apply equally across race, gender, age, disability, ethnicity, socioeconomic status, etc?

6. The operation

Does your algorithm support human decision, rather than override or replace?

7. The maintenance

Are you periodically revisiting your methodology?

Are you updating and retraining your model when new data is introduced?

Are you processing new data and variables with the same inquiry as the original model?

Are the errors you received as expected? Is your algorithm performing on task?

8. The governance

Is the governance group diverse and all inclusive?

Do they have a wide range of skills necessary to understand all the elements of the algorithm?



References and Further Reading

[Algorithm Assessment Report – Summary of Findings & Algorithm Assessment Report](#) – October 2018 - Department of Internal Affairs and StatsNZ – New Zealand

[Principles for safe and effective use of data and analytics](#) – May 2018 – Privacy Commissioner and Government Chief Data Steward – New Zealand

[Initial code of conduct for data-driven health and care technology](#) – September 2018 - Department of Health and Social Care – UK Government

[A Path to Social Licence – Guidelines for Trusted Data Use](#) – August 2017 – Data Futures Partnership NZ

[New Zealand Data and Information Management Principles](#) – August 2011

[The Governance of Decision Making Algorithms Workshop Report](#) – July 2018 – Inspired by a workshop held at the Swiss Re Institute (Centre for Global Dialogue) Rüschlikon (Zürich)

[TRIPOD Checklist: Prediction Model Development and Validation](#)

[Cardiovascular Disease Risk Assessment and Management for Primary Care](#) – 2018 – Ministry of Health

[Population-based Funding Formula \(PBFF\)](#) – 2016 – Ministry of Health

[Virtual Diabetes Register \(VDR\)](#) – 2018 – Ministry of Health

[Can administrative health utilisation data provide an accurate diabetes prevalence estimate for a geographical region?](#) – 2018 – Counties Manukau DHB and Ministry of Health

[Can an Algorithm tell when kids are in danger?](#) – January 2018 – New York Times

[Developing Predictive Risk Models to Support Child Maltreatment Hotline Screening Decisions](#) – March 2017

[Ethics & Algorithms Toolkit](#) – Center for Government Excellence (GovEx) at John Hopkins University, the Civic Analytics Network at Harvard University, the city and county of Dan Francisco and Data Community DC.



The Emerging Health Technology (EHT) team is part of the Ministry of Health Data and Digital Directorate, responsible for understanding and advising on the impacts of new technology across the health and disability system.

EHT are creating Technology Advice and Frameworks to:

- give an introduction to new technologies that are being developed or used in the health sector
- help set the scene for any future conversations had where technology may be applied
- cover where the technology is currently in use
- highlight what impacts it may have on current models of care; and
- present general considerations and/or case studies for health sector stakeholders

Our intended audience is those who are interested in whether emerging technologies will benefit their health deliverables, or who maybe just want a bit more information on what it's all about.

This document is not intended to endorse a specific product or device, but to provide a snapshot of what is happening both locally and/or internationally, and where the major health interest points are.

This is the first step in discovering a technology. There are many other aspects to consider, whether these are funding, technical or clinical, however this is merely to provoke thought, and send you on to authoritative sources.

What do we mean by **Emerging**?

We look at where technology sits within the McKinsey Three Horizons of Growth model. The question we ask is whether this is an improvement to the current model, like improved road tyres, or whether it is disruptive, like driverless cars?

What part of **Health** does this affect?

This could apply to all parts of the health system from clinician to operations, and from population, to primary and specialist care.

How does this relate to **Technology**?

Algorithms relate to all types of technology and they are a key enabler to making technologies more effective.



MANATŪ HAUORA



**Continue the
conversation
in Yammer**



www.yammer.com/emerginghealthtechnology/

