**A new approach to patient stratification**

**Background**

Many PHOs are currently using a risk prediction approach to stratifying patient populations, with the aim of offering targeted services to those most likely to benefit from them. The typical approach is derived from work largely done in the UK in the 2000s, which focusses upon predicting the probability of acute hospital admission as the basis for identifying patients for whom practice teams could provide enhanced services.

The current risk prediction approach has now been used for some time in New Zealand, and it is increasingly clear that it has important limitations. It does not identify mental health need very well, nor is it clear that the blunt event of an acute hospital admission is the best basis on which to target a range of enhanced primary care services. The sheer prediction of the probability of a hospital admission doesn’t necessarily mean that the enhanced services of primary care can prevent that admission, or that a patient can see benefit from those services. There is an emerging view that the time has come for a newer and better approach to stratifying populations and offering enhanced, targeted services.

A better approach should have the characteristics of:

* Identifying patients that have capacity to benefit from a primary care intervention.
* Focussing on services that will reduce inequities in health outcome.
* Being more focussed upon supporting clinical decision making, and providing a wider range of information to support clinicians.

**An approach to developing a better tool**

There are two elements to developing a better tool: firstly characterising the patients and services that clinicians seek to match up, and secondly developing a statistical method for identifying those patients. One approach to the first part would be to form a small group of clinicians to guide the process. They would have the tasks of:

1. Agreeing qualitatively upon the kinds of patients they want to identify and the range of services that those patients would benefit from. This is a clinically led discussion, but will have to be informed by analysts with a specific understanding of available data and the strengths and weaknesses of statistical methods, so that the resulting definitions can be operationalised. The more conditions and types of patients reflect issues of inequity at this stage, then the more the resulting model will have a positive impact on inequitable outcomes.
2. In light of the agreed characteristics of patients that the tool should identify in step one, review a series of patient notes and identify examples of those patients in a real dataset. This will include retrospective information about patients, since the goal is purely to identify the patients one way or another for subsequent analysis. This then becomes the training dataset for statistical analysis.

Once a dataset has been generated that includes clinically identified examples of the patients of interest, the statistical work of generating a model that can identify such patients prospectively can begin. Historically such models have tended to be based upon relatively simple logistic regression techniques, but there is also the capability to investigate methods such as neural network models that are increasingly straightforward to implement and may give improved performance. This statistical process should link with the clinical group as the analysis proceeds, so that the decision on when a model is good enough is informed by both technical and clinical understanding of the model performance.

Issues to consider during the modelling phase include:

* To what extent can wider datasets on social circumstances be made available for the modelling (and subsequently on a routine basis for implementation)?
* To what extent can and should a model be tailored for local circumstances (which may include fundamentally different rates of various kinds of disease, as well as pragmatic considerations such as different local availability of datasets)?

Once a satisfactory model has been generated, it will have to be implemented at PHO and practice level. This will require design of the data flows from PHO back to practices, and design of the resulting presentation so that it provides the most useful feasible information to clinicians. Again, clinical leadership of the design is likely to generate the most effective result.

**Discussion**

The approach outlined here would take the activity of stratifying patient populations for the purpose of offering enhanced services to the next level. It is significantly more clinically led than the historical approach for developing risk prediction tools, but this is likely to mean that it is more effective, particularly in identifying patients who are likely to benefit from services, rather than just patients who are at risk of a poor outcome.

Much of the final state of the existing risk prediction tools depends upon the locally determined state of data collection and data sharing arrangements. The ability to use national datasets (including prescribing, inpatient, outpatient and community mental health data) would cut through such issues, make it far easier to develop a nationally consistent tool, speed the analysis phase of the process, and simplify some aspects of the final implementation. However, it would be dependent upon PHOs having regular access to national datasets with raw NHI, and permission to use those datasets for this purpose. This would have to be the subject of negotiation with the Ministry of Health.

Similarly, information from routine datasets outside health care may well contribute to a more effective tool (for example, being on a waiting list for social housing, or participation in education). There are limits to the feasibility of accessing such data with identifiers, but if any realistic possibilities to access such datasets emerge, they should be pursued.