New care models

Risk stratification: Learning and Impact Study

Operational Research and Evaluation Unit, NHS England

Our values:
clinical engagement, patient involvement,
local ownership, national support

July 2017

For queries please contact:

Operational Research and Evaluation Unit | NHS England

Skipton House, 80 London Road, London, SE1 6LH

England.ORET@nhs.net

Executive summary

Introduction

Risk stratification, or predictive modelling, is used to predict future adverse events, such as unplanned hospital admissions, which are costly, undesirable, and potentially preventable. It is used by healthcare systems in two ways:

- To 'case find' those individuals at risk of such an adverse event, who might have this risk mitigated by being offered a proactive intervention; and
- For population health planning, to understand the distribution and health needs and experiences of different cohorts of such patients.

It forms a core part of the Multispeciality Community Provider (MCP), Primary and Acute Care System (PACS), and Enhanced Health in Care Homes (EHCH) frameworks.

This is the second 'learning and impact study' from the vanguard programme. It draws on findings from five case study visits to MCP and PACS vanguards, a brief literature review, and consultation with senior stakeholders. It has been prepared by NHS England analysts seeking to learn lessons from these vanguards' experiences. It is targeted primarily at those looking to introduce new care models – such as system leaders and transformers aiming to improve local or regional healthcare systems, and clinicians and local analysts seeking to learn technical lessons from the vanguards. It should also be read by central policy makers to draw lessons from across the new care models for future policy; and by the new care models programme team aiming to understand certain aspects of the vanguards in more depth.

What benefits can risk stratification bring?

If effective, risk stratification, used as part of a wider care model, can have a range of benefits:

- Case finding can ensure that individuals at risk of an adverse event can be offered an intervention designed to reduce that risk. This means **that it could support a reduction in a wide range of 'triple fail' outcomes** (such as emergency admissions, premature nursing home admissions, falls, or bed sores) that are simultaneously high cost, low quality, and represent a poor patient experience.
- In stratifying a population by risk, it can also be used as a way **to identify and target appropriate proactive interventions**. It can ensure that the highest-risk patients receive appropriate care for their needs (such as the input of multi-disciplinary teams); medium-risk patients might be referred to a lighter touch intervention (such as social prescribing); and lower-risk patients could be managed through usual care and self-care.
- It can also be used as a population health planning tool, **enabling commissioners and providers to gain a detailed picture of the future risk profile of its population**, allowing them to design care pathways and target funds and interventions appropriately.

What works, and what doesn't, in implementing risk stratification?

The five vanguards studied in this report have successfully implemented the 'building blocks' of an effective risk stratification process. Most are using predictive modelling; implementation is clinically-led and embedded within local pathways; and there is a clear link to proactive, multi-disciplinary team-based care for high-risk individuals. However, there is scope for these models to become more sophisticated, especially if the aim of reducing emergency admissions is to be realised. Six key areas for development, based on the best available published evidence supported by the research undertaken here, are outlined below.

1. **Use predictive models.** There are three types of risk stratification: clinical judgement; threshold modelling; and predictive modelling. The first two have been shown to be of limited effectiveness – there is limited research evidence to suggest that clinical judgement accurately predicts risk of admission (Allaudeen 2011), and threshold modelling is highly susceptible to regression to the mean (Roland & Abel2012). Of the five vanguards we visited, four were using predictive models.

2. **Consider outcomes beyond emergency admissions, and work with lower-risk individuals.** All vanguards were predicting the risk of emergency admission (usually in the next 12 months), and all were focussing on the highest risk strata of population. However, focussing risk stratification only on the highest-risk individuals is likely to limit its overall potential impact. This is because the majority of admissions are in the top 20%, not top 2%, and not all of these admissions will be preventable (Roland & Abel 2012). There is also scope to look at a broader set of 'triple fail' outcomes (beside unplanned admissions), some of which might be more amenable to reduction.

3. **Think about impactability.** Some high-risk patients are more likely to have their risk lowered by a given intervention than others. Whilst most vanguards recognised this concept (called "impactability"), few are addressing this phenomenon in any formal way, other than through clinical judgement. There are a range of approaches to improving impactibility – including gap analysis and a focusing on ambulatory care sensitive conditions. Doing so could increase overall efficacy and cost-effectiveness, and (with some approaches) help reduce health inequalities (Lewis 2011).

4. **Offer a suite of proactive care approaches.** All five vanguards we visited are referring risk-stratified patients to multi-disciplinary teams (MDTs). However, recent evidence reviews suggest that whilst MDT working may improve quality of life and patient satisfaction, it shows limited impact in reducing unplanned hospital care. Interventions such as remote monitoring, and additional clinical support to care homes had greater evidence (Nuffield Trust 2017). The evidence base is currently equivocal, and emerging, and findings for MDTs may improve. Nevertheless, there is scope for vanguards to look at a broader set of proactive interventions, particularly if extending provision from the top 2% of risk to the next 20% or so, where lower-cost, lighter touch interventions might be needed.

5. **Consider information governance as early as possible.** Information governance (IG) for risk stratification can be complex, particularly since reforms to the law in 2012 denied CCGs a legal basis to hold patient-identifiable data. All vanguards interviewed noted that IG was a key challenge in implementing risk stratification, particularly where they wanted to broaden the data they included to look at social care and community data. Vanguards also

noted that IG requirements were often unclear, and were interpreted differently by different individuals and organisations, making it difficult to develop a coherent plan to address these challenges.

6. **Implement risk stratification from the bottom up, with local, clinical leadership.** Risk stratification has been implemented before in the NHS, most notably through more centrally-led policies such as Direct Enhanced Services. Vanguards in this study emphasised the importance of a clinically and locally led approach to implementation, which built on existing CCG approaches and structures. In particular, they drew a contrast between the ability to get buy-in and engagement from a project that was locally-led, compared to one that seemed 'imposed' from the Department of Health or NHS England.

What are the implications for others thinking about taking up risk stratification?

This report describes how a selection of vanguards is using risk stratification in ways that are broadly consistent with the national policy direction (as set out in the New Care Model frameworks). There are numerous ways in which their approach diverges from best practice as suggested by the evidence base. There is also currently limited conclusive evidence linking the use of risk stratification with a reduction in emergency admissions in the vanguards visited, although further evaluations are planned. This is also in line with the findings from the wider literature, which suggests that risk stratification will not have a significant impact on non-elective admissions but can provide value in case finding and population health planning. There are therefore opportunities for developing more sophisticated approaches including by looking at a wider set of outcomes and risk strata, using impactability models, and by offering a more diverse set of preventive interventions that are then continually refined through feedback loops in the data.

1 Purpose of this document

This is the second 'learning and impact study' from the vanguard programme. It draws on findings from: five case study visits to MCP and PACS vanguards; a brief literature review; and consultation with senior stakeholders. It has been prepared by NHS England analysts seeking to learn lessons from the vanguards' experience. It should be read primarily by those looking to adopt new care models including:

- System leaders and transformers, to improve local or regional healthcare systems;
- Clinicians and local analysts, to learn technical lessons from the vanguards;
- Central policy makers, to draw lessons from across the new care models for future policy; and by
- New Care Models programme teams, to understand certain aspects of the vanguards in more depth.

It is a document in two main parts which, while mutually reinforcing, can be read separately to enable more focused use of time. The first half reviews the available literature on the topic while the second half presents findings from some light touch qualitative research with vanguards. The method we used for the work is set out in Annex 1 with a bibliography in Annex 4 and full vanguard summaries in Annex 5. Annexes 2 and 3 provide detailed background on the concepts of positive predictive value & sensitivity.

2 Overview of risk stratification

This section presents an overview of the literature on risk stratification. It draws out the key concepts that should be considered when deciding on an approach to risk stratification.

2.1 Introduction to risk stratification

Risk stratification (often called predictive modelling or predictive risk modelling) is used to predict future adverse events for individuals. These events, such as unplanned hospital admissions, are costly, undesirable, and potentially preventable. Risk stratification is used in two main ways:

1) To 'case find'. To identify those individuals who are at high risk of the predicted event (commonly the top 2% or the next 2-20%), and to offer these patients a proactive intervention designed to reduce their risk of experiencing the adverse event.
2) For population health planning. Here, commissioners and providers use risk stratification to inform their understanding of likely future service demand for different risk cohorts, including their geographic distribution, health needs and co-morbidities, and current pathways of care.

Most commonly, risk stratification has been used to predict the risk of unplanned hospital admissions, usually over a 12 month period. However, it can be used to predict any event that represents a simultaneous failure of all three pillars of the triple aim (i.e. an event that is low quality, gives a poor patient experience, and is high cost) and over different time periods (e.g. 30 days, 6 months, 2 years). Other examples of 'triple fail' events are developing bedsores, falls, losing independence, experiencing an "over-medicalised death[1]", or readmission to hospital. Such events can also be disease-specific, for example predicting which patients with diabetes will experience an episode of diabetic ketoacidosis, or which individuals with kidney disease will start haemodialysis prematurely (Lewis, Kirkham, Duncan & Vaithianathan, 2013).

2.2 Risk stratification in the NHS

Over the last decade or so the Department of Health (DH) and the NHS have promoted risk stratification as a mechanism to: identify patients who will benefit from preventative health interventions; enable better planning of health-related services; and to decrease health-related costs (Curry et al., 2005).

It has been identified as central to three of the New Care Model types – MCPs, PACS, and Care Homes – and is included in all three frameworks, primarily as a means to case find, but also with some emphasis (particularly for PACS) on its use as a resource planning and population health management tool. All frameworks suggest that risk stratification can be used both to support those at highest risk, and those with more on-going care needs/ at a medium or lower risk.

[1] E.g. a death where an individual receives 'life-sustaining treatments, such as mechanical ventilation, that most [patients] indicate they prefer to avoid when faced with less than a year to live (Barnato et al 2009).

2.2.1 Approaches to risk stratification

Historically, there have been three main methods used to risk-stratify and segment a population: clinical perspective, threshold modelling, and predictive modelling:

Clinical perspective

Description: Clinicians use their medical knowledge and training, combined with their knowledge of their patients, to identify individuals at high risk of requiring unplanned healthcare. Clinicians often prefer this method due to a belief (often unfounded) that tools are difficult to use, require information that is not readily available, and have questionable accuracy or utility.
Accuracy: Studies have indicated that clinical perspective is less accurate than predictive models, with one study suggesting that the predictions made by doctors, nurses and case managers of varying seniorities (regarding risk of readmission) were statistically no more accurate than chance (Allaudeen et al 2011). Where clinicians can add significant value is in **identifying which patients are most likely to benefit from an intervention (to be 'impactable')**, and in tailoring interventions for individuals.
Positives: The method is already widely used by clinicians, and is acceptable to them, and may have particular benefit in identifying patients who are 'impactable' (Paton et al 2012).
Negatives: Clinicians' accuracy is limited by three factors, (1) they cannot scan and consider entire populations regularly and repeatedly, (2) they cannot make predictions about patients they do not see, and (3) they are susceptible to cognitive biases. Overall, research evidence indicates **that clinical judgement alone provides limited ability to predict population risk** (Allaudeen et al 2011, Lewis et al 2011).

Threshold modelling

Definition: It came to prominence in the 1980s and 90s, including being used in the prominent Evercare pilots[2]. It is a rules-based, 'catch-all' method that identifies any individual who meets a defined 'high-risk' threshold. An example of such a threshold is: *anyone over the age of 65 years who has had two or more hospital admissions in the previous 12 months*. No statistical modelling is used.
Accuracy: Threshold modelling is more accurate at identifying individuals at *historic* risk, rather than at *future* risk, of an event. As a result, it is heavily susceptible to the phenomenon of "regression to the mean", which states that that if a variable is extreme on its first measurement (e.g. hospital admission rate in the past 12 months), it will tend to be closer to the average ("regress towards the mean") on its second measurement (e.g. hospital admission in the next 12 months) – after one extreme event, the next event is statistically likely to be less extreme. This means that threshold modelling is even less accurate than chance in predicting future risk.
Positives: The method uses readily-accessible data and requires minimal financial and time investments.
Negatives: The model itself is **predisposed to selection bias and regression to the mean**. All patients' use of services varies over time. Threshold modelling is likely to pick up a patient just after their highest use of services (which is when they are about to regress to the population mean). This means that on average patients identified by threshold modelling will have lower rates of unplanned hospital admission in the future even without preventive intervention.

[2] These showed no evidence of impact on reducing emergency admissions (see Gravelle et al 2007).

Predictive modelling

Definition: This method uses multiple regression, neural networks or other forms of artificial intelligence to make predictions of future risk, based on individuals' past characteristics. Inputted data can include: socio-demographic; diagnostic; prior utilisation/cost; pharmacy data; health status/functionality; and clinical data (Curry et al., 2005). Three NHS models have been developed – Patients at Risk of Re-hospitalisation (PARR), the Combined Predictive Model (CPM) and the Patients at Risk of Re-hospitalisation over 30 days (PARR-30). **All are free to use, and have been shown still to be accurate predictors of risk**. Unlike PARR, the CPM can also predict risk for patients who have not had an emergency admission in the last 24 months. Active promotion of these tools ceased in 2011. Since then a range of other tools, often developed by commercial organisations, have been used by the NHS. These are generally associated with a licence cost, so organisations should think carefully about whether they offer any added benefits over the free tools (e.g. increased accuracy, better interfaces/ ease of use), and whether any such benefits offset the cost of procurement.
Accuracy: Varies with the statistical technique used, the data used, and the event being predicted. One study found that predictive risk modelling is about twice as accurate as threshold models at predicting future events (Cousins, Shickle & Bander, 2002). No predictive model is perfectly accurate; therefore it will generate false positive and false negative results in addition to true positive and true negative predictions (Lewis 2015). Clinical judgement is often used to review the output of predictive models, partly for case finding but sometimes with a suggestion that this increases accuracy. There is is as yet no robust evidence that clinical review increases the accuracy of predictions (Allaudeen 2011), however clinical review is a key component of case finding.
Positives: **It is the most accurate risk stratification method available**; it is less susceptible to regression to the mean and to cognitive biases, and can help reduce health care inequalities – though it still has some limitations. Social care and demographic information can be incorporated into the model, in addition to health and healthcare data; and the tools can also be used to predict "triple fail" events relating to social care (e.g. losing independence)
Negatives: Relative to the other two techniques, predictive modelling can be more expensive to implement and typically requires analytical support to use. Commissioning Support Units or private companies are commonly commissioned to provide this support. It also has limitations (as all approaches do) in its ability not just to predict risk , but predict which high-risk patients will be most likely to benefit from the preventive care being offered. This phenomenon ("impactibility") is discussed in more detail below.

2.3 Potential impacts of risk stratification

Risk stratification is most commonly framed as a means to reduce emergency admission rates (and hence costs), by proactively managing the care of the highest need, highest cost users – though, as noted above, it can be used to predict any 'Triple Fail' event. However, whether or not risk stratification actually results in a reduction in emergency admissions is dependent on more than just the accuracy of the model. A range of factors need to be considered:

- Concept 1: The accuracy of the tool within each stratum of risk (the "positive predictive value" of the tool);
- Concept 2: The proportion of emergency admissions (or other "triple fail" event) in each stratum of risk (the "sensitivity" of the tool);
- Concept 3: Whether the emergency admissions due to be experienced by the target population are preventable ("impactibility");

- Concept 4: Whether the proactive interventions put in place will be effective (and cost-effective) at preventing them ("efficacy" and "cost-effectiveness").

These considerations are discussed in more detail below. The focus is on emergency admissions, as this is what all five vanguards are looking at; however, if sites were also to use this approach for other 'Triple Fail' outcomes they would need to be aware that assumptions around relative accuracy and impactability might be different.

2.3.1 Concept 1: The accuracy of predictive tools

Any assessment of the accuracy of a predictive tool needs to take account not just how accurately it correctly identified a high-risk individual, who then went on to have a hospital admission (*true positives*), but also the frequency of *false positives* (people identified as high-risk, but who did not go on to have a hospital admission), and *false negatives* (people identified as low-risk, but actually went on to have a hospital admission). This is important as both false positives and false negatives have potentially negative implications for patients – summarised in the table below.

Table 1.4: Definition and implications of false positives and false negatives

	Definition	Implications
False positive	Patient incorrectly identified as high-risk - *e.g. were predicted to have an emergency admission, but did not*	• Intervention 'wasted' • Needless anxiety for patient; • Possible over-investigation and subsequent over-treatment.
False negative	Patient incorrectly identified as low-risk – *e.g. were predicted not to have an emergency admission, but did have one*	• Interventions not offered that should have been offered; • Possible escalation of symptoms; • Unwarranted reassurance

There are many ways of gauging the accuracy of risk stratification. Two of the most useful measures are sensitivity, and positive predictive value (PPV). These two complementary measures can be traded off against each other. If we lower the risk threshold we use to identify people and increase the number of people identified, we will increase the *sensitivity* (a higher absolute number of people at high risk will be identified) but decrease the *PPV* (the proportion of those identified that are actually high risk will lower). There will be more true positives but also more false positives. This is discussed in more detail in Annexes 2 and 3.

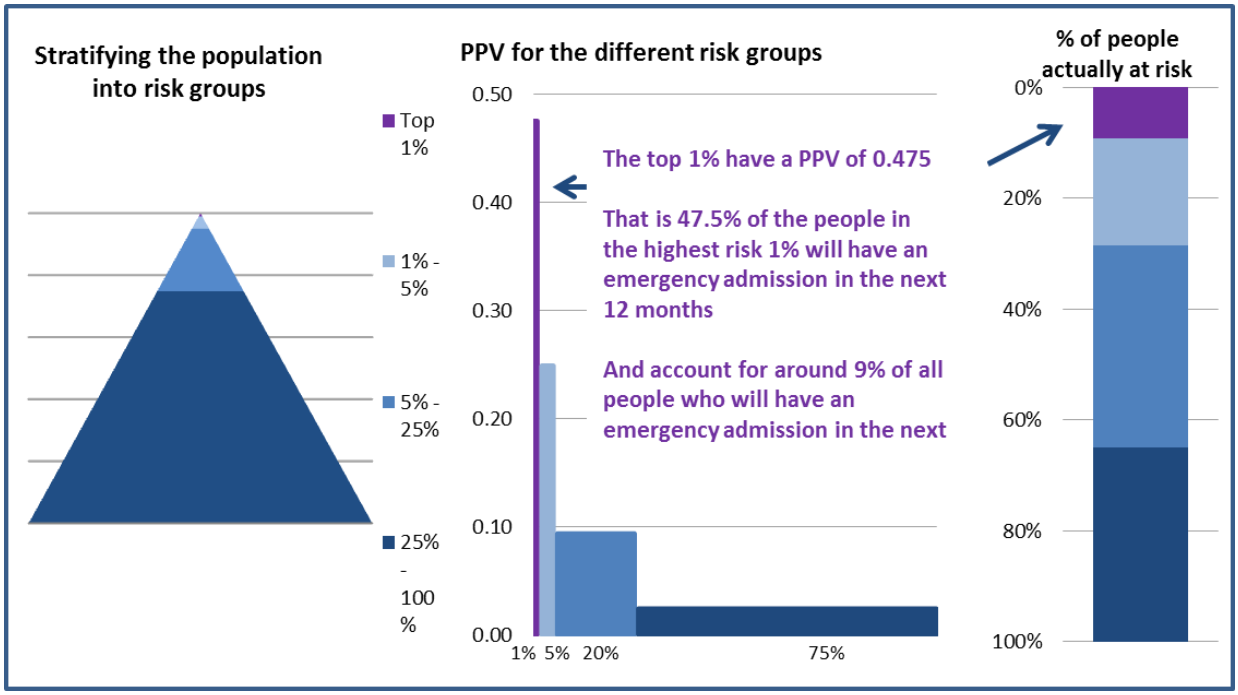
2.3.2 Concept 2: Impact on emergency admissions

Risk stratification programmes, including the five vanguards visited as part of this study, are often designed to offer preventive care to the top 2% or 5% of highest risk individuals within the population. This approach is based on the rationale that these individuals account for a disproportionately high proportion of healthcare costs. However, whilst their use might indeed be disproportionate, they do not account for the majority of admissions across the population as a whole. For example, as detailed below, the top 1% of people identified through the Combined

Predictive Model would typically account for 9% of total admissions; 91% of admissions[3] would come from outside this group.

Given that a higher number of admissions (in absolute rather than percentage terms) come from lower risk patients, the greatest reduction in emergency admissions would come from reducing the risk of a larger segment of the population (such as the top 20%) not just the top 2% (Wallace et al 2016). Figure 1.4 shows a stylised image of how the distribution of risk actually translates into the burden on service use.

Figure 1.4: Burden of future utilisation



2.3.3 Concept 3: Preventability of target population's emergency admissions

Preventability of admissions

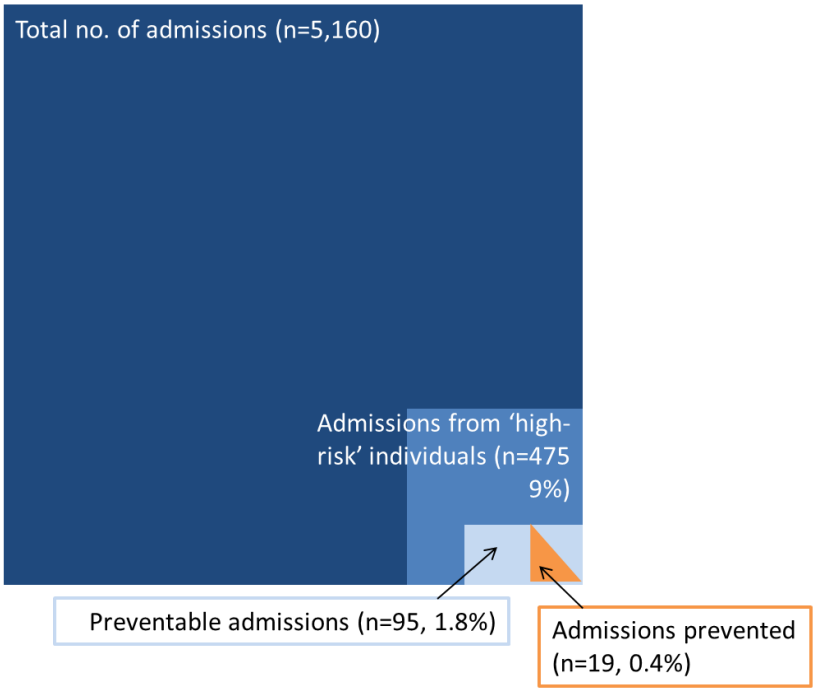
In addition, not all of the target group's emergency admissions will be preventable, and even if they are, one cannot assume that any intervention will be 100% successful in preventing them. For example, applying the revised Combined Model (used in the example above) to a hypothetical local health system of 100,000 population:

- A local health economy with 100,000 population would expect 5,160 patients to have at least one emergency admission a year;
- Risk stratification's top 1% would equate to 1,000 people;
- We would expect 475 (9.2%) 5,160)of these high risk individuals to be admitted at least once in the next year (true positives);
- Thus, 4,685 patients (91% of total admissions) having at least one emergency admission would not have been identified (false negatives);

[3] N.B. data does not distinguish between planned and unplanned admissions.

- In addition, not all of these 475 admissions would be preventable. For example if we assume that 20%[4] of the population at highest risk of admission are impactable – 1.8% (or 95) of all emergency admissions.
- Evidence suggests that around a fifth of these amenable admissions would actually be prevented by current interventions (Tian et al 2012)[5]. **In total, this local health system could expect interventions targeted using risk stratification of its top 1% to prevent 0.4% (or 19) of all emergency admissions (Figure 1.5).**
- Using a unit of £1,600 per admission this would save a total of £30,000, therefore **for an intervention to be financially neutral would have to cost less than £30 per person targeted.**

Figure 1.5: Illustration of preventable admissions from a GP practice of 100, 000 people (based on CPM model)



This same modelling approach can be applied to different segments of the population. If the proportion of the population targeted is increased (say, to 20%), then there will be a higher number of true positives, a reduced number of false negatives (admissions not identified), and a greater number of admissions will be prevented. However, this will also cost considerably more, as interventions need to be offered to a greater number of individuals. For example, using the same model and population as above, but **applied to the top 20% of the population, 2.5% (or 116) emergency admissions would be prevented.** Using a unit of £1,600 per admission this would save a total of £190,000, therefore for an intervention to be financially neutral **would have to cost less than £9.50 per person targeted (see Annex 3 for full workings).**

[4] Based on assumptions around the distribution of ambulatory care sensitive conditions.
[5] Assuming all local authorities performed to the level of the best-performing local authority

Impactability of populations

The success of risk stratification depends not just on identifying those most at risk of an adverse event, but rather in identifying those who are most at risk *and* most likely to respond to a given intervention – to be 'impactable'. There are three main forms of impactability modelling. All should improve the efficiency of models (by targeting people more effectively), but whereas the first two of these approaches will help reduce healthcare inequalities and should be promoted, the third may worsen healthcare inequalities and must therefore be avoided (Lewis 2010):

1. *Giving priority to patients with diseases that are particularly amenable to preventative care.* There are a range of methods to doing this, the most common being to focus on the subgroup of high-risk people who have one or more ambulatory care-sensitive (ACS) conditions. This is a list of conditions where there is evidence to suggest that if they are managed optimally in the community then they should not result in an unplanned admission. Such conditions are more prevalent in more deprived areas so targeting support to high-risk patients with one or more ACS conditions should help reduce inequalities.
2. *Giving priority to patients with 'gaps' in their care* (e.g. a patient with ischemic heart disease not taking an antiplatelet drug). Such gaps are found more frequently in more deprived areas (the "Inverse Care Law") so targeting support at high-risk patients with a high gap score should help reduce inequalities.
3. *Excluding patients who are least likely to respond to preventative care.* In the United States, high-risk patients are sometimes systematically excluded from preventive care based on clinical or other individual characteristics (such as previous non-compliance, low activation, or language barriers). These methods are at least controversial, if not illegal in the UK, as they disadvantage particular groups, and therefore widen health inequalities.

Ultimately, the aim of impactibility modelling is to identify the form of preventative care best matched to a high-risk patient's individual characteristics. Here, the best approach for a particular patient is modelled and refined based on feedback loops consisting of demographic, diagnostic, and social characteristics. Models might include both varying the means of contact (nurse or doctor; time of intervention), the content and frequency of the intervention (for example, based on cultural preferences), and the use of incentives.

2.3.4 Concept 4: Efficacy of proactive interventions

Risk stratification's impact is tightly linked to the efficacy of the interventions or follow up support offered to risk-stratified patients. Even if an admission is technically 'preventable', an appropriate intervention or care plan will need to be put in place to do this.

There are a wide range of interventions and approaches designed to reduce risk of hospital admission for complex patients (see, for example, Imison et al. 2017). In vanguards, these include provision of support from multidisciplinary teams, extensivist care services, remote monitoring, self-management support and social prescribing. We do not intend to rehearse the evidence for each intervention here. Instead, we outline some of the factors that should be considered when thinking about which interventions to offer, and the likely impact on outcomes. Vanguards should also consider the extent to which a given intervention is *acceptable* and *deliverable* in their local area; even if an intervention is technically effective, if it is poorly implemented minimal impact will be seen.

Efficacy, and cost-effectiveness, of interventions

Interventions to reduce hospital attendance vary in the strength of evidence behind them (see Imison et al 2017 for a full summary, and section 4.2 of this report for a discussion of Vanguards' approaches). In addition, providers also need to be aware of the risks of supply-induced demand, disbenefits, and the varying cost of different interventions.

Supply-induced demand refers to the phenomenon whereby a new service can result in an increase (rather than the intended reduction) in demand. This is cited as one reason why interventions to improve the quality and coordination of care tend to show increases in quality, but no reductions in cost (Roland & Abel 2012).

Interventions may also have disbenefits – both in absolute terms (for example, a trial of a COPD self-management programme was terminated early due to higher mortality in the intervention group) and in relative terms (there is relatively robust evidence that self-management can widen health inequalities due to inequitable access) (Roland & Abel 2012; Health Foundation 2014). These considerations should be factored in when assessing the potential outcomes of risk stratification.

Lastly, the costs of interventions should be considered; in particular, there is a need to trade-off the proportion of the population targeted (a bigger population = more scope for impact) against the potential cost of interventions (a bigger population = greater cost of interventions) and the efficacy of the intervention (see Billings et al 2012 for a breakdown of costs of re-admission by risk band).

What does the evidence mean for the vanguards and those adopting the new care models?

Conduct a range of impact assessments

Information governance arrangements can be complex, so it is important to conduct a Privacy Impact Assessment before introducing a new risk stratification programme. An equality impact assessment is essential to ensure that the health care inequalities are not inadvertently exacerbated. And an ethical review is important to ensure that the benefits of the programme outweigh the harms caused by false positive and false negative predictions.

Use predictive models

Predictive models are the best-evidenced method for predicting individuals' risk. Threshold modelling is heavily affected by regression to the mean, and clinical perspective is subject to bias and may be no better than chance.

Think about 'impactability', not just absolute risk

Impactability tools, including gap analysis and a focus on ambulatory care sensitive conditions can help identify those high-risk individuals who are most likely to have their risk lowered by a given intervention. It is important to use an approach that helps reduce health care inequalities and to build feedback loops to improve the impactibility model over time.

Try a suite of options to manage risk stratified patients' care – different interventions will be effective for different individuals

There are many different interventions offered to high-risk patients (MDT support, self-management support, remote monitoring, etc.). The evidence on the efficacy of these interventions varies, with self-management support particularly well evidenced but variable evidence for MDTs. All need further work to understand what works for whom, and why.

Recognise the distribution of risk across the population

Risk stratification is unlikely to have a profound impact on a population's emergency admissions rate if it is limited to the highest-risk individuals. This is because the majority of admissions come from outside the top 1-2% of risk. In addition to targeting the top risk strata, vanguards may want to consider offering progressively lower cost preventative care to the top 0.5%-5% and the top 5%-20% risk strata.

3 Overview of risk stratification in the vanguards

This section provides a brief overview of risk stratification in each of the five vanguards in this report, highlighting the type of predictive model and data used, and how risk stratification fits with other vanguard interventions. Full descriptions of each site can be found in Annex 5.

3.1 Erewash

Type of model: Threshold, with some clinical input
Use of model: Case finding

Erewash's risk stratification model is used to determine the case load of care-coordinators (CCs). These CCs manually review any patient over 18 who have been seen in A&E, MIU, OOH, and look through primary care records on SystemOne to identify persons over 65 years who are using primary and/or acute services frequently OR persons over 80 years old who have not accessed a health professional in the past 12 months. The records of these individuals are then passed onto MDT teams, who discuss the individual and plan proactive care.

3.2 Morecambe Bay

Type of model: Predictive modelling, with clinical input
Use of model: Case finding

Morecambe Bay started implementing their predictive modelling risk stratification tool in December 2016. The tool – enhanced Combined Predictive Model (developed by the King's Fund, but tailored to the local environment) – is run by Midlands and Lancashire CSU. It is a multiple regression model that uses SUS, GP practice data, co-morbidities and demographic data to predict the likelihood that each registered patient will have a non-elective admission in the next 12 months. The findings appear on GPs' EMIS system, with GPs expected to use their clinical judgement to decide whether or not to refer onto local Integrated Care Communities.

3.3 North East Hampshire and Farnham

Type of model: Predictive modelling, with clinical input
Use of model: case finding and population health planning

NEHF currently use the ACG predictive modelling tool – though they plan to move onto the IPA (a tool developed by the CSU, based on Johns Hopkins' ACG algorithms) in the next few months. The ACG tool uses a mixture of logistic and linear regression. It was initially implemented as part of a Direct Enhanced Service (DES), and is now being relaunched. It uses around 250 predictive factors (including primary and secondary care utilisation and demographic data) to predict risk of hospitalisation in the forthcoming year, and risk of being 'high cost' in the forthcoming year. GPs have access to a 'dashboard' presenting results from the risk stratification, alongside primary and secondary care activity and (in the future) community, social care, and mental health data. GPs can then refer the identified patients to community MDTs. Typically, the model looks at the top 2% or 5% of the population, though NEHF are exploring varying this.

3.4 Sunderland

Type of model: Predictive modelling, with clinical input
Use of model: Case finding

Sunderland uses the Q-admissions predictive modelling tool, developed by ClinRisk LTD. The tool uses around 30 risk factors recorded on primary care systems to identify the top 2 or 3% of the population at risk of avoidable admission within the next 12 months. All members of the MDT are then able to add individuals who they think might benefit from an MDT approach (this might be based on wider information sources such as hospital discharge summaries, social care data, wider clinical indicators (e.g. on the frailty list) and soft intelligence from team members). However, the extent to which this happens varies across the city.

3.5 Tower Hamlets

Type of model: Predictive modelling, with clinical input
Use of model: Case finding, population health planning and to support the commissioning of integrated care services.

Tower Hamlet uses the Q-admissoins tool stratifying the population by their risk of hospital admission over the forthcoming year. The tool draws on linked hospital data (SUS) and GP practice data under a section 251 exemption, deanonymisation can only be done by general practice clinicians and is cumbersome- patient by patient. The tool was originally used to identify people for a 'community virtual ward' type offer, but it was found that many of the patients in the high risk cohort had combined physical and mental health issues so appropriate services for these patients were also provided. Limits to the approach have been identified, including the sped of the data flows limiting its day-to-day use, as well as the volatility of the risk scores meaning use for long-term planning and budgeting was difficult. The team is also keen to examine the possibilities of using data outside of the routinely collected health data sets to predict risk.

3.6 Summary table of risk stratification methods in vanguards

	Method	Analytical tool	Outcome predicted	Accuracy	Data used
Erewash	Threshold model	None	None, not a predictive model. Selects individuals > 65 years who use primary and/ or acute services frequently, and >80 years who have not accessed a health professional in last 12 months	N/A threshold model, accuracy not known (but would be low)	Primary care data, acute admissions
Morecambe Bay	Multiple regression predictive model	Combined Predictive Model[6], which has been enhanced by Midlands and Lancashire CSU	Risk of an unplanned hospital admission in the next 12 months	PPV[7] for emergency admissions for the original CPM was 30% for the top 3.5% of patients at the highest risk. The model has now been updated though.	Primary and acute data, supplemented with demographic information
NEHF	Multiple regression predictive model	ACG tool[8], developed by Johns Hopkins	Risk of an unplanned hospital admission in the next 12 months. This is to be expanded to other variables in the future.	Accuracy not known.	Primary and acute data. Currently working on adding mental health and social care data.

[6] https://www.kingsfund.org.uk/sites/files/kf/field/field_document/PARR-combined-predictive-model-final-report-dec06.pdf
[7] Positive Predictive Value (PPV) is a reflection of the number of patients who actually had an emergency admission in the year following prediction out of all the patients who were predicted to have an emergency admission. So if 300 of the 1,000 patients predicted to have an ED, had an ED, the PPV would be 30%. It is usually expressed for a particular subset of the population (e.g. the 10% or 3% at most risk of admission). It's accuracy increases the greater the risk of admission – so should be higher for the top 3% than top 10%.
[8] https://www.hopkinsacg.org/

Sunderland	Multiple regression predictive model	Q-admissions[9], developed by ClinRisk LTD	Risk of unplanned hospital admission in the next 12 months.	PPV for emergency admissions for the top 10% of patients at highest risk was 42%, sensitivity of 39% (Cox & Coupland 2013)	Primary care data, supplemented with a variety of other information
Tower Hamlets	Multiple regression predictive model	Q-admissions[10]	Risk of an unplanned hospital admission in next 12 months	PPV for emergency admissions for the top 10% of patients at highest risk was 42%, sensitivity of 39% (Cox & Coupland 2013)	Primary care and SUS data, supplemented by clinical judgement.

[9] http://www.qadmissions.org/
[10] http://www.qadmissions.org/

4 Thematic analysis of vanguards' approaches to risk stratification

This section analyses the findings from the qualitative research, drawing out common approaches to designing and implementing risk stratification across the five vanguard sites. It describes vanguards' approaches in relation to case finding, case management and population health planning. It then presents an overview of the barriers and enablers for risk stratification, and the early evidence of its impact.

4.1 Case finding: What tools are the vanguards using, and how are they using them?

The risk stratification process is broader than simply the risk stratification method used, vanguards vary in:

- Which tool they use;
- What data sources feed into the risk stratification model (what variables are they including and why);
- What outcome they predict (what variables they are predicting and why);
- Whether and how clinicians and other sources of intelligence are involved in the process;
- What strata of their population they target.

Each of these issues is discussed in turn below.

4.1.1 Most of the vanguards in this report are using predictive modelling

Risk stratification models range from the simple (either based on GP views, or using threshold models) to the more complex (predictive models using regression or decision trees). Four of the five vanguards are using relatively sophisticated regression models (though the particular model varies); Erewash is using a manual threshold model.

There is diversity in the tools chosen, with four different approaches being used. Mostly, this reflects the fact that vanguards have been working closely with CCGs to build on previous work/ previously commissioned tools (e.g. in Sunderland) – suggesting that the tools are likely to be the 'best fit' for local environments.

4.1.2 Vanguards' risk stratification tools use a variety of input data

The predictive power of risk stratification tools is heavily dependent on the type and quality of the input data. Broadly, the data sources used can be split into six categories: socio-demographic; diagnostic; prior utilisation; pharmacy; health status/ functionality; and clinical data. All five of the vanguards in this report, in common with risk stratification across England, are using data on prior primary and secondary care service utilisation, alongside demographic factors. Some models are beginning to look more broadly at potential predictors. For example, Sunderland is meeting with social care to understand whether their risk stratification metrics

(more than 15 hours of planned care, or the use of the alarm more than 50 times[11]) could be usefully added to their tool.

However, there was also a recognition from the five vanguards that there needed to be a degree of pragmatism in limiting the number of variables included, particularly where they might be time consuming, or require IG permissions, to acquire. For example, Tower Hamlets noted that there might be value in adding more social factors – such as social isolation or living in crowded housing – but that there was a need to be pragmatic and work within the confines of existing data flows.

This use of multiple categories of data is encouraging, and fits with the wider evidence base. This suggests that predictive accuracy is increased by: more categories of variables, as the variables from each category have poor predictive power on their own; the use of diagnostic and utilisation data together with demographic data; and the use of pharmacy data as a reasonable substitute for diagnostic data (Curry 2005).

	Socio-demographic	Diagnostic	Prior utilisation/cost	Pharmacy data	Health status/functionality	Clinical data
Morecambe Bay	Y	Y	Y			Y
NEHF	Y	Y	Y	Y	Y	Y
Sunderland	Y	Y	Y	Y	Y	Y
Tower Hamlets	Y	Y	Y	Y	Y	Y

4.1.3 All vanguards are predicting risk of unplanned hospital admission in the next 12 months

Currently, all five vanguards in this study are predicting the risk of individuals requiring an unplanned hospital admission in the next 12 months, with only NEHF and Tower Hamlets currently planning to expand this to include other outcome measures.

NEHF's ACG tool does this both by taking a wider view of hospitalisation (looking at planned and unplanned admissions, and bed days), and also looking at variables related to cost. NEHF are also considering broadening this even further to look at PROMS, wellbeing, and clinical indicators. Given the issues with using risk stratification to reduce emergency admissions (see 2.3.1), and the fact that the evidence for some of the interventions put in place suggests that they might be more effective at reducing length of stay, rather than admissions (Imison et al 2017), this might be a sensible approach.

4.1.4 Vanguards are supplementing the data included in their tool with other data sources and clinical judgement

As noted above, most vanguards' risk stratification processes still include some degree of clinical input, as well as additional data sources. Once a list has been generated, most of the vanguards display this alongside additional social, psychological, and clinical information, and

[11] Inclusion of additional metrics should be done with care; some metrics are more susceptible to regression to the mean (such as the use of the alarm on many occasions).

rely on GPs or MDTs to consider this additional information, alongside their own clinical judgement, to further refine the list.

For example, in Sunderland an initial list is generated by the risk stratification tool. This is then supplemented with data from other sources including looking for: households with multiple individuals on the list (e.g. to understand where you might have at-risk carers); those identified as living in a care home or with dementia; those with complex medication regimes; and soft intelligence from GPs (*'if we are talking about them, we are worried about them'*), pharmacy staff and receptionists.

Similarly, Morecambe Bay has introduced findings from their Frailty Project. This is a coding system that sits within EMIS and SystemOne, and codes frail individuals based on number of co-morbidities. This categorisation is then visible to a GP when a patient's risk stratification profile is bought up.

Overall, the use of additional sources of data has positive and negative aspects. Positively, it shows that vanguards are effectively integrating risk stratification to make use of existing data sources, expertise and intelligence. Clinical judgement also might contribute to the selection of 'impactable' patients (see below). On the other hand, there might be concerns that the involvement of clinicians and MDTs in altering and refining lists might introduce biases (e.g. Lewis 2011).

4.1.5 The target population and impactability

The success of risk stratification depends on identifying those who are most at risk *and* most likely to respond to a given intervention – to be 'impactable'. Typically, risk stratification interventions have targeted the top 2 or 3% of their populations, due to these individuals being responsible for a disproportionate number of admissions (though not, as noted in section 2.3.1, the majority of admissions), with some evidence that hospital-avoidance interventions are most effective for this cohort (Krause, 2005; Peikes et al 2009). However, the vanguards noted that the ability of interventions to reduce admissions might vary according to whether:

- The health of individuals within the highest needs cohort can be influenced by an intervention – or whether their health needs are too complex even for a relatively intensive MDT or extensivist support model (Curry et al., 2005);
- The few individuals within this cohort can be accurately identified (Lewis et al., 2011);
- Whether the 2-3% stratum is appropriate in all localities. In particular, both Sunderland and NEHF noted that, given the significant variation in health inequalities and health burden across their vanguards this proportion will need to vary. For example, Sunderland and NEHF noted that in some of their localities the top 2% will need input from the integrated care team, whereas in others it will be the top 4%. NEHF are now extending this work further to give a target proportion for each locality.

The five vanguards in this study are yet to go beyond recognising the issue of impactability to try and model or identify those individuals who might be most likely to benefit from their interventions – despite the existence of (some) impactability models developed to assist this (see Lewis 2010). Similarly, there was minimal focus on tailoring specific interventions for

particular target groups. This may reflect the development stage of their care models, and/or the lack of evidence in the area[12].

4.2 Case management: what are vanguards doing for the people they identify?

Whilst risk stratification has some value in improving clinicians' understanding of their population's risk factors, its impact on emergency admissions and other outcomes measures is mostly dependent on the efficacy of the interventions that follow it. All five vanguards are using risk stratification to identify individuals to be referred onto multi-disciplinary community teams – in various forms. NEHF's MDT model includes specialist palliative care input. Erewash's model is slightly lighter-touch, using a 'care coordinator' model, where one individual, as opposed to a multidisciplinary team, coordinates care for a person – though they can refer onto an MDT as appropriate. For all, risk stratification is clearly embedded within a wider pathway of care, much of which has been developed as part of the vanguard.

It is notable that all vanguards have settled on a similar model of care, and that all have gone further than simply giving the risk stratification outputs to GPs, and expecting them to act on it. This reflects a recognition by all vanguards that GPs have neither the time, nor the knowledge and skills to act alone to offer the holistic support needed to reduce risk amongst the population. This recognition comes both from previous experience implementing GP-led risk stratification (particularly in NEHF), as well as the wider evidence base (and NHSE guidance) which suggests that proactive case management should involve input from specialists from a range of disciplines, in an MDT setting (NHS England 2016a, 2016b, 2016c; American Case Management Association 2013).

However, the evidence on the efficacy of community-based MDTs in reducing hospital admission is still relatively weak. Indeed, in the Nuffield Trust's recent review of approaches to shifting the balance of care case management and care coordination presented only mixed evidence, particularly on cost reduction. One systematic review's (Stokes et al 2015) subgroup analysis showed a small, non-significant reduction in short-term secondary care utilisation following community MDTs, suggesting it is an intervention worth investigating further, but not that it is necessarily effective.

Conversely, the Nuffield Trust's review found that remote monitoring, community-based EoL care (which may be covered by MDTs), and additional clinical support to nursing and care homes all had relatively robust, positive evidence (Imison et al 2017).

4.3 Population health planning: how are vanguards using risk stratification to understand their populations?

Risk stratification can also be used as a tool to support CCGs in population health planning. Of the five vanguards interviewed, only two – NEHF and Morecambe Bay – had done any population modelling work, and that work was relatively limited at this stage. This was despite

[12] For example, the most recent Cochrane reivew on effective interventions for people with multimorbidity did not look at effectiveness for particular cohorts, and showed generally weak evidence of efficacy (Smith, Wallace & O'Dowd 2016).

most CCGs being heavily involved in procuring and running the risk stratification tools. There is therefore scope for CCGs to be encouraged to make further use of this.

4.4 What works (and doesn't) with implementing risk stratification in vanguards?

This section draws on findings from the qualitative analysis to identify individual and institutional factors which support or hindered the implementation of risk stratification.

4.4.1 'Fit' with the local health care system policies

All five vanguards were working closely with their CCGs to develop their risk stratification policies. In some cases, they were building on and extending previous CCG work, in others, working with the CCG to procure and implement new processes. Where this was effective, vanguards reported that it meant they built on, rather than duplicated, existing work; learnt lessons from previous attempts; ensured that leadership was local; and made full use of the management and analytical support a CCG can offer. For example:

- Erewash have expanded the CCG's previous case identification process;
- Morecambe Bay are working with the CCG's newly procured analytical tool, and using CCG business and analytical support;
- Similarly, Sunderland use the CCG's case identification system, and are provided with business and analytical support, and;
- Tower Hamlets are building on a process developed locally through the Pioneer programme.

There were varying views on the role of the vanguard in accelerating or shaping the risk stratification process. Some interviewees felt that the vanguard had accelerated the rate of change, and improved buy-in, whilst others felt it was something the local health economy would have been doing anyway. Some vanguards also reported that linking with CCGs meant they could build on the increased awareness of, and 'comfort' with, data and analytics developed as a result of using risk stratification. Given that this research did not speak to non-vanguards about their development of risk stratification approaches, it is hard to draw firm conclusions on the additional benefit offered by the programme.

Situating risk stratification within the wider local health and social care economy also improved links to, and efficacy of, follow-on interventions. Indeed, in several cases, the integrated care teams that risk stratified patients are referred onto were developed alongside risk stratification, as part of the vanguard, and are mutually reinforcing. Interviewees also noted that embedding risk stratification within the wider system meant that the purpose of the tool was clear to stakeholders; there was a clear link to improved patient care, which improved engagement.

4.4.2 Clinical ownership

Clinicians and particularly clinical leadership, was both a barrier and enabler for vanguards. NEHF and Sunderland both reported that strong clinical leadership and ownership of risk stratification was key to effective implementation. In particular, both noted the importance of those who will use the tools designing the overall process; not only does this help ensure that tools are fit for purpose, and fit with existing processes, but it ensures local teams have

ownership over the model. One vanguard highlighted how it helped convince staff that risk stratification was more than 'just a list' churned out by a computer, that the guidance produced was clinically-driven, and clinically authored.

NEHF, in particular, drew a contrast with their previous attempt to implement risk stratification, which was a top-down process, instigated in response to an NHSE DES. It was reported that this stimulus limited the buy-in locally at the time; few clinicians understood why risk stratification was being used, or agreed that the way it was being done was the right way. This time, with clinical leadership, implementation is progressing more smoothly.

Indeed, several vanguards noted that clinicians could continue to be a barrier to implementation. Some clinicians remained sceptical of the accuracy and usefulness of the predictive tools. A common concern was that 'the tools identify patients who are already on our radar'. This lack of confidence contributed to a feeling that risk stratification was a 'tick box' exercise that diverted resources away from patient care, to identifying the top 2% of the population - without any effective services to refer them on to.

Sometimes GPs also simply lacked the time to review, consider, and then implement risk stratification. For example, Morecambe Bay reported that their GPs were already over-burdened, and Sunderland that, despite significant resource being dedicated to staff engagement, awareness of the guidance remained relatively low amongst staff (including partner organisations).

4.4.3 Data sharing and information governance

Information governance and the 2012 changes to the Data Protection changes

In 2012 data protection changes were introduced that prevented CCGs processing patient identifiable data. This had a significant impact on the development of risk stratification tools, many of which were under development (as was the case for Erewash). In 2013, permission was granted for the data processing of personal data for risk stratification purposes, but it meant that patient data needed to go through a process of 'pseudonomisation' whereby an alternative ID number is applied to patient data before it can be processed[13].

Organisations now have to make special applications in order to be able to process data for risk stratification[14]. Organisations need to be aware both of the relevant IG procedures, and the costs of processing requirements, and factor this into the design of any risk stratification model.

All vanguards were at different stages in progressing their data sharing agreements, but all reported that this had been a major barrier to implementation. All noted that, should these issues not be resolved, the accuracy and effectiveness of the tools would be significantly diminished. For example:

- Erewash reported that their failure to procure a predictive risk stratification tools is a function of the 2012 reforms, which have meant that they are unable to share data.

[13] Full details are available here: https://www.england.nhs.uk/ourwork/tsd/ig/risk-stratification/
[14] A full list of approved organisations is available here: https://www.england.nhs.uk/publication/list-of-risk-stratification-approved-organisations/

They do not feature on the list of organisations which are cleared to carry out risk stratification. They are currently being supported by NHSE to address this.

- NEHF and Sunderland are hoping to link social and community care data into their risk stratification tool. This would be used not just to feed into the predictive model, but as part of a wider 'dashboard' of care indicators produced for each patient to allow ICTs to understand their needs. However, IG and data sharing issues are currently delaying this.

Several vanguards noted that the issue with IG wasn't that they were unable to share data per se, but that the appetite for risk varied considerably across different organisations and individuals. They noted that a clear steer from NHSE and NHS Digital, highlighting the importance of risk stratification, encouraging action to overcome barriers, and setting clear parameters around IG, would be very helpful.

4.5 Early evidence of impact

There is minimal robust evidence provided thus far about the impact that adopting risk stratification approaches has had on patients' outcomes. All vanguards are at an early stage of implementing and evaluating their risk stratification tools, therefore findings should be interpreted with care – we would be unlikely to see much impact on system-wide metrics at this stage. One might expect early anecdotal findings of impacts on patient experience and satisfaction, or GP/MDT confidence in using the tool but, for the most part, local evaluators are yet to report on these findings[15].

Of the limited evidence that has emerged:

- There is anecdotal evidence from NEHF that including risk stratification information on ambulances' care planning records has helped paramedics to tailor their support to the particular patient visited.
- There is not yet any evidence of a decline in unplanned hospital admissions in NEHF and Sunderland (the two areas with the best-developed risk stratification). As noted in section 2.3.1, this might reflect the fact that this is an insensitive choice of outcome measure, rather than suggesting risk stratification is ineffective per se. It also reflects the fact that risk stratification has not been in place long.
- Sunderland is reporting emerging evidence of vanguard-wide reductions in their length of stay and delayed transfer of care measures. Qualitative findings from local evaluations will help clarify the extent to which this can be attributed to risk stratification.

[15] This report was finalised in early 2017, local evaluations reported in March/ April 2017.

What can other sites learn from these Vanguards?

Ensure local clinicians are involved in designing and implementing risk stratification

All Vanguards highlighted the importance of clinical buy-in and clinical leadership. NEHF, in particular, highlighted the contrast in effectiveness between the NHSE-mandated approach of 2012 (via DESs), and the clinician-led approach this time. They felt GPs would be more much more likely to actively engage with risk stratification this time, and to embed it sustainably within their usual practice.

Recognise that IG may be an issue

All Vanguards raised IG as an issue in implementing risk stratification, particularly when extending input data to social care. NHS organisations are subject to fair processing guidelines and may need to access a data 'safe haven' in order to implement risk stratification. Organisations implementing risk stratification will need to allocate appropriate time and resources in order to ensure that they can process data in line with current guidance. This may involve processing taking place outside the organisation.

Risk stratification data should not be presented alone

Most vanguards presented risk stratification findings to GPs as part of a dashboard of measures. This dashboard often also included community and social care data, allowing GPs to make a more holistic, informed decision about an individual's risk and impactability..

Embed risk stratification within wider pathways

Risk stratification's efficacy is heavily linked to the effectiveness of the interventions onto which patients are referred, and the extent to which community interventions are available. Ensuring that these interventions are part of a clear, well integrated pathway that clinicians are aware of, and buy into, is essential to ensuring risk stratification data is not just seen, but used.

Recognise that MDT working is unlikely to reduce overall emergency admissions on its own and a range of interventions need to be considered

All vanguards are referring the patients identified via risk stratification (and deemed to be suitable through clinical judgement) onto MDTs. Recent evidence reviews (Imison et al 2017) suggest that whilst case management (via an MDT or other method) may improve quality of life and patient satisfaction, it has limited impact in reducing unplanned hospital care. There is therefore scope for vanguards to look more broadly at other interventions, such as remote monitoring, and self-management support. There is also scope for vanguards to consider which tier of risk they are targeting for intervention. MDT working and case management may be effective for some individuals, but should be part of a suite of measures, and realistic expectations should be in place before it is commenced. Qualitative objectives and measures around length of stay may be better measures of the impact of risk stratification than admission frequency.

Questions for STPs and commissioners to consider when commissioning and implementing risk stratification

- **Tool selection**: What are the capabilities of the tool being used? Does it support a predictive model? Is it suitable for both the needs of the organisation? For example, at what level is the population segmentation viewable - is it at the practice level or practice group level, or Is it broader? Is this suitable for the context in which is being implemented?
- **Information governance:** How thoroughly have information governance issues been addressed by the organisation processing the data and the organisations reviewing the results? Are data processing agreements in place and has approval been gained? Has an impact analysis and equalities impact assessment been carried out?
- **Case finding and population planning**: Who will be reviewing the outputs of the risk tool and with what intended goals? What resources are available for onward referral? How will the people identified be supported, and how effective are the interventions being offered?

Annex 1: A note on the method

A note on the method used. This study examines a selection of the PACS and MCP vanguards. As such, a set of criteria was used to select which vanguards to examine in depth. This was guided by risk stratification information contained in vanguards' logic models, and gathered by account managers during structured conversations in June 2016. Two PACS (Morecambe Bay and North East Hampshire and Farnham) and three MCP (Erewash, Sunderland, and Tower Hamlets) vanguards accepted the invitation to take part in the impact study. Together, the five vanguards represent the range of approaches to, and processes for, risk stratification across the new care models programme. Data were collected in face-to-face group discussions with the five individual vanguards in late 2016. Attendees were nominated by the vanguards, though we suggested representation from the programme, the associated CCG, analytics, and primary care. Notes taken by NHS England analysts were checked for accuracy by those representing vanguards. Additionally, specific questions about the analytical tools used for risk stratification were answered (via email) by those who utilise the tools. Where the report refers to 'all' or 'most' vanguards it is referring to those five who were consulted – unless explicitly stated otherwise.

Annex 2: Sensitivity and Positive Predictive Value

In gauging the accuracy of a tool, two complementary measures are often used: sensitivity, and positive predictive value (PPV).

Sensitivity is the proportion of all high-risk individuals that the model correctly identifies as high-risk:

$$Sensitivity = \frac{true\ positives}{true\ positives + false\ negatives}$$

PPV is the proportion of those identified as high risk by the model, that are actually high risk:

$$Positive\ Predictive\ Value = \frac{true\ positives}{true\ positives + false\ positives}$$

Importantly, there is a trade-off between sensitivity and PPV. If we lower the risk threshold we use to identify people and increase the number of people identified, we will increase the *sensitivity* (a higher absolute number of people at high risk will be identified) but decrease the *PPV* (the proportion of those identified that are actually high risk will lower). There will be more true positives but also more false positives. For example: Results presented for the model with Inpatient, A&E, Out Patient and GP data[16]

% of population	PPV	Sensitivity
0.5%*	0.562	0.048
1%	0.475	0.092
2.5% *	0.378	0.176
5%	0.294	0.285
10% *	0.22	0.422
20%**	0.145	0.6
25% *	0.133	0.649

Notes: * values for these Strata are approximate. They actually apply to 2.4%, 9.9% and 25.18%. ** interpolated from other strata presented in the paper..

This example explicitly shows the trade off in selecting the different risk strata. Selecting the top 0.5% of risk scores will deliver a high PPV (around 56% of the people identified will go on and have an emergency admission in the next 12 months) but a low sensitivity (these people will only be around 5% of the total number of people to have an emergency admission in the next 12 months).

While targeting the top 20% of risk scores will identify a higher absolute number of true positives and higher sensitivity (identifying around 60% of all people who go on to have an emergency admission) but also a far higher number of false positives and lower PPV (only around 15% of those identified will have an emergency admission).

[16] Billings J, Georghiou T, Blunt I, et al. Choosing a model to predict hospital admission: an observational study of new variants of predictive models for case finding.

Annex 3: Applying CPM to the top 20% of the population

The revised Combined Model can be applied to a hypothetical local health system of 100,000 population, but looking at the top 20%. In this case:

- A local health economy with 100,000 population would expect 5,160 patients to have at least one emergency admission a year;
- Risk stratification's top 20% would equate to 20,000 people;
- We would expect 2,900 of these high risk individuals to be admitted at least once in the next year (true positives);
- Thus, 2,260 patients (40% of total admissions) having at least one emergency admission would not have been identified (false negatives);
- In addition, not all of these 2,900 admissions would be preventable. For example if we assume that 20%[17] of the population at highest risk of admission are impactable – 12% (or 580) of all emergency admissions.
- Evidence suggests that around a fifth of these amenable admissions would actually be prevented by current interventions (Tian et al 2012)[18]. **In total, this local health system could expect interventions targeted using risk stratification of its top 20% to prevent 2.4% (or 116) of all emergency admissions.**
- Using a unit of £1,600 per admission this would save a total of £190,000, therefore for an intervention to be financially neutral would have to cost less than £9.50 per person targeted.

[17] Based on assumptions around the distribution of ambulatory care sensitive conditions.
[18] Assuming all local authorities performed to the level of the best-performing local authority

Annex 4: Bibliography

Key reading

Lewis, Curry & Bardsley (2011) Choosing a predictive risk model: a guide for commissioners in England, *Nuffield Trust* – discusses the issues in selecting a risk model and implementing in practice.

Lewis (2015) Next steps for Risk Stratification in the NHS – outlines the current NHS England learning on risk stratification and its benefits and potential limitations.

Roland & Abel (2012) Reducing emergency admissions: are we on the right track? *BMJ* – discusses the issues in linking risk stratification to emergency admissions reduction.

Allaudeen et al (2011) Inability of providers to predict unplanned readmissions, *J gen Intern Med*, 26 (7)

Barnato et al (2009). Development and validation of hospital "end-of-life" treatment inten- sity measures. Med Care. 47(10):1104.

Billings et al (2012) Development of a predictive model to identify inpatients at risk of readmission within 30 days of discharge (PARR-30), *BMJ Open 2(4)*

Cox & Coupland (2013) Predicting risk of emergency admission to hospital using primary care data: derivation and validation of QAdmissions score, *BMJ Open* 3(8)

Curry N, Billings J, Darin B, Dixon J, Williams M, Wennberg, D. 2005. Predictive Risk Project: Literature review.

Imison, Curry, Holder et al (2017) Shifting the balance of care: Great expectations, *Nuffield Trust*

Lewis (2010) 'Impactibility models': identifying the subgroup of high-risk patients most amenable to hospital-avoidance programs, *Millbank Quarterly 88(2)*

Lewis (2015) Next steps for Risk Stratification in the NHS

Lewis, Curry & Bardsley (2011) Choosing a predictive risk model: a guide for commissioners in England, *Nuffield Trust*

Gravelle et al (2007) Impact of case management (Evercare) on frail elderly patients: controlled before and after analysis of quantitative outcome data, *BMJ, 334:31*

NHS England (2016a) The multispecialty community provider (MCP) emerging care model and contract framework

NHS England (2016b) Integrated primary and acute care systems (PACS) – Describing the care model and the business model.

NHS England (2016c) Enhanced Service Specification - Avoiding unplanned admissions: proactive case finding and patient review for vulnerable people 2016/17.

Paton, Wilson & Wright (2012) Predictive validity of tools used to assess the risk of unplanned admissions: A rapid review of the evidence, *Centre for Reviews and Dissemination*

Roland & Abel (2012) Reducing emergency admissions: are we on the right track? *BMJ*

Smith, Wallace, O'Dowd & Fortin (2016) Interventions for improving outcomes in patients with multimorbidity in primary care and community settings, *Cochrane Database of Systematic Reviews*

Stokes J, Panagioti M, Alam R, Checkland K, Cheraghi-Sohi S, Bower P. Effectiveness of case management for "at risk" patients in primary care: a systematic review and meta-analysis. *PLoS One 2015; 10*

Tian, Dixon, Gao (2012) Data briefing: emergency hospital admissions for ambulatory care-sensitive conditions.

Wallace, Smith, Fahey & Martin (2016) Reducing emergency admissions through community-based interventions, *BMJ; 352*

Annex 5: Detailed summaries of vanguard approaches to risk stratification

Erewash

Aim of risk stratification: The aim is twofold. First, patients aged 65+ years who have a high number of acute and primary care visits are identified so that proactive care can be planned and provided. Second, patients aged 80+ years that have not accessed a health professional in the past 12 months are identified to make sure they are receiving appropriate care services.

Leadership of risk stratification: The day-to-day work is both CCG- and vanguard- led. The CCG have leadership of the overall risk stratification process.

Description of risk stratification processes: Care Co-ordinators (CCs), who are administrators that physically sit within GP practices, manually review primary care records on SystemOne to identify the frequent- and non- users of primary and acute services. The CCs pass the details of the identified individuals to MDTs weekly. The MDTs discuss the individuals and plan their proactive care. Existing health and social care services are utilised.

Role of vanguard in risk stratification processes: CC positions were introduced by the CCG in 2012 – prior to the vanguard. However, vanguard funding has been used to finance more CC positions. The vanguard is involved in the CC's case finding, only. The vanguard is not involved in the care planning during the MDT meetings. The vanguard has therefore increased the existing risk stratification processes, but has not altered them in any way.

Early signs of impact of risk stratification processes: Anecdotal evidence suggests CCs are significantly reducing GPs' workload, enabling GPs to spend more time with patients. Currently, no other impact is measurable, as there is no feedback loop between the CCs, GPs, MDT, analytics and the CCG.

Distinguishing feature of risk stratification in the vanguard: The vanguard is not currently using an analytical tool. The Erewash region used a tool previously, but the 2012 health reforms resulted in information governance (IG) issues that blocked data-sharing. The use of the tool subsequently ceased, and the tool was no longer commissioned.

Future goals for risk stratification processes: There is a willingness to better understand risk stratification and what it can offer the vanguard.

Morecambe Bay

Aim of risk stratification: The primary aims are to identify individuals who are most likely to require a non-elective hospital admission in the next 12 months, and to plan and provide proactive care for these individuals. A future aim for the CCG is to use risk stratification information to plan and commission services.

Leadership of risk stratification: The overall process will be CCG-led. The vanguard's integrated care communities (ICCs) and GPs will lead the day-to-day risk stratification work.

Description of risk stratification processes: Processes have not been fully developed or agreed due to the recent roll-out of the analytical tool. However, the CCG and vanguard have established a general protocol. Midlands and Lancashire CSU will run the analytical tool and update and maintain the risk stratification data. The CSU will also provide analytical support to GPs, for example assist in interpreting the tool's output. All GPs in the vanguard will have access to risk stratification output, and the onus will be on the GPs to refer to, and use, the data as they see fit.

GPs will supply the ICCs with the details of individuals at high risk of a non-elective hospital admission in the next 12 months. The ICCs will subsequently plan care for these individuals.

The resources and structure are not consistent across the vanguard and therefore case identification and case management will probably vary across practices and ICCs.

Role of vanguard in risk stratification processes: It will be up to the GPs within the vanguard to access the risk stratification data, and to use the data. The vanguard-developed ICCs will provide case management.

Early signs of impact of risk stratification processes: It is too early for impact to be measured, and there is currently no formal assessment plan. However, processes and impact will be discussed at monthly ICC meetings (that GP leads also attend).

Distinguishing feature of risk stratification in the vanguard: The organisation of the vanguard around Integrated Care Communities (ICCs) means that there is a good structure and good resources in place to support risk stratification.

Future goals for risk stratification processes: As the new risk stratification processes become 'business as usual', the vanguard will look to align the processes across the ICC teams. They are looking to include ambulance services data and Patient Activation Measure data. There is an

aspiration to focus on patients with moderate (rather than high) risk, in line with the widespread acknowledgement that those with moderate risk might benefit most from an intervention.

North East Hampshire and Farnham

Aim of risk stratification: There are two aims. First, risk stratification information will be used to monitor and shape services and thereby inform commissioning. Second, it will be used to identify the individuals with the highest 2% risk of an unplanned hospital admission in the next 12 months.

Leadership of risk stratification: It is being led by the vanguard, specifically by local GPs within the vanguard.

Description of risk stratification processes: The five integrated care teams (ICTs) within the vanguard will each use the analytical tool to identify the 2% of their respective populations at most risk of an unplanned hospital admission. These individuals will then be discussed and their care planned at ICT MDT meetings.

Role of vanguard in risk stratification processes: The vanguard leads and undertakes all risk stratification processes.

Early signs of impact of risk stratification processes: There has been some levelling out of increases in A&E attendances and emergency admissions, though it is too soon to establish either whether this is significant, or if it is attributable to risk stratification. The vanguard are conscious of the limitations of hospital admissions as an outcome measure, and are investigating other options.

Distinguishing feature of risk stratification in the vanguard: Risk stratification is being led by local GPs. As part of the vanguard GPs have protected time to carry out, and reflect on, risk stratification. NEHF, together with the CSU, are building their own risk stratification tool, using the algorithm from the Johns Hopkins ACG tool.

Future goals for risk stratification processes: The vanguard are investigating changing the 'outcome' measure of the predictive tool. They are looking at measures such as multi-morbidity, polypharmacy, or emergency bed days. They would also like to establish using the tool to monitor and check the performance of services.

Sunderland

Aim of risk stratification: The primary aim is to identify the 3% of the adult population at greatest risk of an unplanned hospital admission. Secondary aims include improving patient experience, improving return on investment, increasing cost savings, and identifying previously unknown unmet need.

Leadership of risk stratification: The QA tool was procured as part of the national DES, with the vanguard workstreams leading and developing the risk stratification process.

Description of risk stratification processes: Q-admissions is run every second month. Its output is supplemented with information from other sources to create a final list of patients at high risk of requiring unplanned health care. The list is taken to one of the five locality-based community integrated teams (CITs). The MDTs within the CIT then evaluate the patients' care needs and develop appropriate care plans.

The CITs are encouraged to be creative in identifying their 3%, and as a result each CIT has different approach.

Role of vanguard in risk stratification processes: The vanguard developed the CITs, which perform case identification and case management.

Early signs of impact of risk stratification processes: There has been no reduction in unplanned hospital admissions at vanguard level, but there has been a reduction in admissions for the 3% of the population identified by risk stratification. The vanguard are aware of the possibility that this reduction is due to regression to the mean and are further investigating the data.

These data have raised questions within the vanguard. Does the data suggest that the risk stratification process is not identifying the right group? What other factors which could be impacting upon the admission rate?

Distinguishing feature of risk stratification in the vanguard: The vanguard has developed a guidance document for risk stratification. The document was co-produced by staff who implement risk stratification. The guidance encourages CITs to be creative in undertaking risk stratification. They hope this will enable the development of effective processes.

Future goals for risk stratification processes: The vanguard would like to look at the outcomes for risk stratified patients and assess the impacts of the risk stratification processes. They will then use this information to align processes across the CITs. The vanguard would also like to ddevelop more sophisticated analytical tools, for example by incorporating the electronic frailty index at population level. Additionally, the CCG would like to use segmentation across the whole population, and to use a greater number of data sources. All of these goals require new or amended information governance agreements.

Tower Hamlets

Aim of risk stratification: to provide appropriate, disease-based care for patients with LTCs. The tool was originally used to identify people for a 'community virtual ward' type offer, but looking at the types of patients in the high risk cohort the team also found a significant number with combined physical and mental health issues and so also worked on a strong RAID offer and redesign of relationships between CMHTs and general practice.

Leadership of risk stratification: The CCG has overall leadership of risk stratification working with the local GP networks.

Description of risk stratification processes: Tower Hamlets uses the Q-admissions tool and stratifies the population by their risk of hospital admission over the forthcoming year. The tool draws on linked hospital data (SUS) and GP practice data under a section 251 exemption. Deanonymisation can only be done by general practice clinicians and is thought to be quite cumbersome- patient by patient. There is then a meeting with practice based clinicians and representatives from the local integrated care teams where they will discuss those in the high risk mandatory category plus a range of discretionary people who can be added. The key factor for who goes on the final list for receiving a more proactive, holistic package of care is not considered to be a patient's risk of admission to hospital but instead is a an assessment of "who would benefit from an holistic care input?" There have been numerous challenges with this approach, discussed below.

Role of vanguard in risk stratification processes: projects to integrate services have a comparatively long history in Tower Hamlets. The vanguard is building on previous work done by the Pioneer project, for example. Risk stratification has been an approach used throughout this process and therefore the vanguard team has continued this work.

Early signs of impact of risk stratification processes: There is limited evidence of impact available although we expect this to be picked up in future evaluation work.

Distinguishing feature of risk stratification in the vanguard: The local team has explored a range of approaches and identified a number of issues with these. The team reports that using the tool for patient identification for service delivery proved cumbersome and data probably

flowed too slowly to be of real use (monthly reporting but slow, cumbersome deanonymisation). The team has also found that admission risk scores are volatile at the individual patient level with high levels of regression to the mean. They note that, across a year, only 20% of these who were high risk at the start of year are still high risk at the end of the year. This has therefore proved not to be an effective way of identifying populations for capitated budgets or for the proactive holistic enhanced GP care packages to be delivered over years with an emphasis on continuity of care and comprehensive geriatric assessment/ monitoring of possible iatrogenic harms. Currently the patients eligible for this practice are clinically defined- mandatory patients are on the palliative care register, have dementia and or are in care homes. There is also therefore an allocated 'discretionary' caseload size per practice so that this service can also be offered to other patients felt to need this additional support- especially those with complex multimorbidity, often with functional impairments, sometimes as 'step down' from the more focussed short term offer from community teams.

Future goals for risk stratification processes: The risk stratification approach is currently heavily reliant on routinely collected healthcare data. However the team has undertaken analysis which suggests that factors like social isolation, functional impairment and overall levels of a person and their social network's resilience, resource and personal knowledge, skills and confidence to self-manage would probably be better predictive variables. The team has been collecting functional and social network data and has put in a request for a251 exemption to combine health and LA data sets (eg property gazetteer, electoral roll) for planning.