Next Steps for Risk Stratification in the NHS
This paper discusses potential next steps for risk stratification within the NHS and has been produced to be used as an additional resource for the Case Finding and Risk Stratification LTC handbook. It will be accessed by a url and will be located in the Risk Stratification part of the handbook.

The resource paper is linked to the Case Finding & Risk Stratification handbook which is part of a set of three handbooks that include MDT & Personalised Care & Support Planning.

Dr Geraint Lewis FRCP FFPH
Chief Data Officer
NHS England
england.cdo@nhs.net

This is a controlled document. Whilst this document may be printed, the electronic version posted on the intranet is the controlled copy. Any printed copies of this document are not controlled. As a controlled document, this document should not be saved onto local or network drives but should always be accessed from the intranet.
Document Title: Next Steps for Risk Stratification in the NHS

Version number: 1.0

First published: January 2015

Prepared by: Dr Geraint H. Lewis FRCP FFPH; Chief Data Officer, NHS England

Classification: Discussion paper

NHS England Equality Statement

"Equality and diversity are at the heart of NHS England’s values. Throughout the development of the policies and processes cited in this document, we have given due regard to the need to:

- Reduce health inequalities in access and outcomes of healthcare services integrate services where this might reduce health inequalities
- Eliminate discrimination, harassment and victimisation
- Advance equality of opportunity and foster good relations between people who share a relevant protected characteristic (as cited in under the Equality Act 2010) and those who do not share it."
Contents

Executive Summary .................................................................................................................. 5

Introduction .............................................................................................................................. 7

2. Using Risk Stratification to address the challenges ..................................................... 7

3. The promise of risk stratification .................................................................................... 8

4. The problems with risk stratification ............................................................................. 8

5. Impactibility ...................................................................................................................... 11

6. Evidence base for preventive programmes ................................................................... 13

The Way Forward ..................................................................................................................... 14

7. Safeguards for NHS organisations .................................................................................. 14

Conclusion ............................................................................................................................... 16

References ............................................................................................................................... 17
Executive Summary

Risk stratification offers the potential to improve the quality and experience of care for patients whilst reducing costs for the taxpayer. However, it is also beset by a number of potential problems. Specifically:

1. The predictive accuracy of many risk stratification tools is modest. No risk stratification tool is ever completely accurate; therefore it is important to consider the potential adverse impact of false positive and false negative results as well as the benefits of true positive and true negative results. By varying the cut-off used to define different strata of risk, NHS organisations can increase or decrease the number of false positives and false negatives. For a risk stratification programme to be effective, the benefits to the population must outweigh the costs.

2. Some of the strategies used to improve the impact of risk stratification programmes could potentially worsen health care inequalities. Some true positive patients may still experience an unplanned hospital admission despite the best efforts of the preventive intervention. By identifying the subgroup of true positive patients who will actually benefit from the preventive intervention being offered, impactibility models can improve the efficiency of a risk stratification programme. Some types of impactibility models can also help reduce health care inequalities; however, other types of impactibility model may worsen health inequalities and must therefore be avoided.

3. Many of the preventive interventions offered in risk stratification programmes appear to increase total costs rather than reduce them. Given the lack of robust evidence to support many of the hospital-avoidance interventions being offered to high-risk patients, there is a pressing need for further research and evaluation.

Any NHS organisation interested in beginning or expanding its risk stratification programme should start by conducting an opportunity analysis. This process involves analysing population data to identify the incidence of low-quality, high-cost, poor-experience events that might be amenable to preventative care. The next step is to conduct an ethical review, perhaps based on the framework published by the World Health Organisation.

Risk stratification programmes should be evaluated using a valid comparator group. A pre-post study does not constitute a valid comparator group; therefore CCGs should consider using other evaluation methods, such as pragmatic randomised controlled trials, propensity score matched cohort studies, or regression discontinuity analyses.
The data generated in any risk stratification programme should be used in a feedback loop to improve the performance of the programme. One of the issues that can hamper both evaluations and feedback loops is the problem of small numbers so NHS organisations should consider working collaboratively with each other and pool their data for analysis.
Introduction

1 Health care systems across the developed world are currently facing a similar set of challenges. Our populations are ageing and chronic illnesses are becoming more prevalent; patients’ hopes and expectations are rising; and budgets are becoming increasingly tight.\footnote{Given these demanding circumstances, policymakers are naturally attracted to any initiative that promises to improve the quality and experience of care while simultaneously reducing overall costs. Hence the popularity of
- payment reforms that promote capitated budgets and payment for performance;
- programmes to promote shared decision-making and better self-management of chronic conditions; and
- initiatives that encourage patients to make greater use of primary care.\footnote{Another strategy that many health care systems are deploying to address these challenges is to use risk stratification as a way of improving the targeting of preventive care.}}

Another strategy that many health care systems are deploying to address these challenges is to use risk stratification as a way of improving the targeting of preventive care.

2. Using Risk Stratification to address the challenges

2.1 In any population, a relatively small number of patients accounts for a disproportionately large fraction of health care costs. In England, for example, roughly half of all hospital bed-days are attributable to just five per cent of the population.\footnote{If these high-cost individuals could be identified early and offered better support and preventive care then it might be possible to improve their health outcomes and their experiences of care while, at the same time, making large net savings for the health service as a whole from averted complications and avoided hospital admissions ‘downstream’.\footnote{Such an approach relies on the ability to identify appropriate patients. We know that there is a rapid turnover of the individuals that make up the highest-risk cohort;\footnote{Therefore, any preventive care intervention must be offered to those people who are at high risk of becoming high-cost in the future – not simply offering preventive care to those people who have been high cost in the recent past. For this reason, risk stratification tools are specifically designed to identify those individuals who are at high risk of experiencing a future adverse event, such as a readmission within 30 days or an unplanned hospital admission in the next 12 months.}}}

2.2 Such an approach relies on the ability to identify appropriate patients. We know that there is a rapid turnover of the individuals that make up the highest-risk cohort;\footnote{Therefore, any preventive care intervention must be offered to those people who are at high risk of becoming high-cost in the future – not simply offering preventive care to those people who have been high cost in the recent past. For this reason, risk stratification tools are specifically designed to identify those individuals who are at high risk of experiencing a future adverse event, such as a readmission within 30 days or an unplanned hospital admission in the next 12 months.} therefore, any preventive care intervention must be offered to those people who are at high risk of becoming high-cost in the future – not simply offering preventive care to those people who have been high cost in the recent past. For this reason, risk stratification tools are specifically designed to identify those individuals who are at high risk of experiencing a future adverse event, such as a readmission within 30 days or an unplanned hospital admission in the next 12 months.

2.3 However, risk stratification is no \textit{deus ex machina}: as we shall see, the predictive accuracy of these tools is modest; some of the strategies designed to improve their impact could potentially worsen health care inequalities; and many of the preventive interventions offered according their predictions actually appear to drive up costs rather than reduce them. In this paper we therefore begin by reviewing the potential
of risk stratification to improve population health before examining each of these three pitfalls in turn. We end by offering some advice to NHS organisations on how to proceed cautiously with risk stratification. Risk stratification offers huge potential to improve the quality, equity and efficiency of care but the NHS needs to minimise the possibility for inadvertent waste or harm.

3. The promise of risk stratification

3.1 Risk stratification aims to identify patients who are at high risk of an adverse event so that these people can be offered preventive care today aimed at averting costly, unpleasant health problems tomorrow. In this sense, risk stratification is analogous to population screening: its aim is to identify people who are more likely to be helped than harmed if they are offered further tests or treatment. As we shall see, no screening test or risk stratification tool is ever completely accurate; therefore it is important to consider the adverse impact of false positive and false negative results as well as the benefits of true positive and true negative results. The aim of an effective risk stratification programme is to ensure that the benefits to the population outweigh the costs.

3.2 Currently, much of the focus with risk stratification programmes in the NHS is on predicting unplanned hospital admissions in the next 12 months. Such admissions are important for three reasons. First, they may be an indicator of suboptimal care; second, they are generally unpleasant and undesirable for patients and their families; and third, they are costly to the health service. There are, however, many other events that meet this “Triple Fail” definition, including readmissions to hospital within 30 days of discharge and admission to a nursing home in the next 12 months. Indeed, risk stratification tools have already been developed using NHS data to predict such events so it will be important for local NHS organisations to consider the potential role of these tools as another way of improving the health of their local population.

4. The problems with risk stratification

4.1 As we have seen, risk stratification holds great promise; however, it is also beset by a range of potential problems. These difficulties include the relatively weak predictive accuracy of many risk stratification tools; the ethics of adjunct tools called impactibility models, which may be used to improve the efficiency of risk stratification; and the lack of evidence for the effectiveness of many preventive programmes that are offered to high-risk patients.

4.2 Predictive accuracy

There are two principal alternatives to predictive models that can be used for predicting a “Triple Fail” event such as an unplanned hospital admission. These alternatives are threshold models and clinical judgement. See Box 1
Threatened models are simple, rules-based criteria. An example of such a rule is to classify as “high-risk” any patients in the population aged 65 or over who have experienced two or more unplanned admissions in the previous year. The advantage of this approach is that it is simple and intuitive. Unfortunately, however, it is undermined by a statistical phenomenon called regression to the mean, which holds that patients who experienced an extremely high frequency of hospital admissions in one year will tend to have fewer admissions the following year, even without intervention. Patients identified using threshold models are therefore likely to experience a reduction in admission rates without preventive care, making it difficult or impossible for a preventive programme to achieve additional benefits for patients who are identified in this way.

The other main option is to ask professionals, such as doctors and nurses to select patients based on their clinical experience. Unfortunately, clinicians are subject to a range of cognitive biases. For example, the availability bias suggests that clinicians are more likely to identify patients that come immediately to mind rather than taking equal account of all patients, including those who have predominantly had contact with other parts of the health service. Ultimately, the evidence suggests that clinicians are less accurate than risk stratification tools at predicting risk. Indeed, a study by Allaudeen and colleagues found that the predictions made by junior doctors, senior doctors, nurses and case managers were statistically no different from chance.

4.2.1 Given the problems associated with the two alternatives set out in Box 1, predictive risk models are currently regarded as the most accurate way to identify patients at risk. However, no predictive model is perfectly accurate. Indeed, a 2011 study by Kansagara and colleagues found that many of the predictive models used for forecasting readmissions to hospital performed poorly. As with any form of screening test, there are four potential outcomes for any individual whose data are risk stratified:

1. **True Positive** (person is correctly identified as being at risk)
2. **True Negative** (person is correctly identified as not being at risk)
3. **False Positive** (person is wrongly identified as being at risk)
4. **False Negative** (person is wrongly identified as not being at risk)
4.2.2 Statisticians use various metrics to describe the accuracy of a screening test or risk stratification tool.\(^{14}\) For example, Kansagara and colleagues defined a risk stratification tool as performing poorly if its \(c\)-statistic was below 0.7. The \(c\)-statistic is an aggregate number that reflects the distribution of true positives and true negatives across all risk scores.\(^{i}\)

4.2.3 In reality, preventive interventions are only offered to individuals certain strata of risk (for example, a Clinical Commissioning Group [CCG] using the PARR tool might choose to offer an intervention to all patients with a risk score of 70 or above).\(^{15}\) For this reason, rather than considering the performance of the tool across all risk scores using the \(c\)-statistic, it is generally preferable to consider the accuracy of the tool only for patients in risk stratum of interest (in this case 70-100). The most useful measures of a tool’s accuracy within a particular risk stratum are the \textit{sensitivity} and the \textit{positive predictive value}.\(^{15}\) See Box 2

**Box 2: Predictive accuracy within risk strata**

The most useful measures of predictive accuracy within a risk stratum are:

- **Sensitivity**, which is the proportion of true positives who are correctly identified as high risk
- **Positive Predictive Value**, which is the proportion of people identified as high risk who are truly positive

4.2.4 Take as an example a CCG with a registered population of 130,000. The CCG decides to offer a preventive intervention to every patient with a PARR risk score of 70 and above.\(^{ii}\) A CCG of this size would typically have 130 such individuals. The sensitivity of the PARR tool for the 70-100 risk stratum is 17.8 per cent and the positive predictive value for this stratum is 77.4 per cent.\(^{15}\) Of the 130 “high risk” people identified by the tool with a risk score between 70 and 100, there would be 100 individuals who would, without intervention, experience an unplanned hospital admission the following year (true positives), while 30 of the 130 would not be so admitted (false positives). Of the remaining 129,870 people in the population, 129,270 people would not experience an unplanned hospital admission the following year (true negatives) while 600 or so people would be wrongly classified as low risk (false negatives).

\(^{i}\) Technically, the \(c\)-statistic is the probability that a randomly selected patient with a future admission will receive a higher risk score than a randomly selected patient who will not have a future admission.

\(^{ii}\) For information governance reasons, patients with a risk score just below this cut-off should be reviewed by a clinician.
4.3 It is important to remember that there are potential harms associated with both false positive and false negative results. The problem with false positive results is that the individuals concerned are offered an intervention to prevent an event that they were not actually going experience. As a result, the preventive intervention would be “wasted” on these people and the resources would have been better spent elsewhere. Moreover, these individuals might experience needless anxiety from being wrongly told that they were high risk, and they might also be subject to over-investigation or over-treatment. For example, such a patient might have their medications reviewed as part of the package of preventive care. As a result, they might be offered more aggressive treatment in an attempt to keep them healthy, treatment that could result in unnecessary side effects. Alternatively, they might be subjected to more invasive tests and investigations because the clinician erred on the side of caution, knowing that the patient had been classified as being at “high risk” of unplanned admission.

4.4 In contrast, the difficulties associated with false negative results are related to unwarranted reassurance. For example, a clinician might downplay the significance of new symptoms because they thought that the patient was “low risk”. As a result, the patient might experience a delay in the detection of an illness and then, when the problem was detected, the illness might be at a more advanced stage – meaning that more invasive treatment might be required, which could be less successful, have more side effects, and be more costly.

4.5 Choosing a cut-off risk score
Clearly, the overall impact of any risk stratification programme depends on the relative frequency, costs and benefits of true positives, true negatives, false positives, and false negatives. An important advantage of risk stratification tools over other prediction methods is that it is possible to trade off the sensitivity of the tool against its positive predictive value. In other words, it is possible to increase or decrease the number of false positives at the expense of decreasing or increasing the number of false negatives, respectively. For example, by choosing a lower threshold (e.g. offering the intervention to every person with a risk score of 60+ rather than 70+), a CCG can reduce the number of false negatives, but at the expense of increasing the number of false positives. In contrast, by choosing a higher threshold (e.g. risk scores of 80+) the CCG can reduce the number of false positive results but in doing so will increase the number of false negatives.

5. Impactibility
5.1 As we have seen, CCGs are able to vary the number of true positives identified by lowering the risk score threshold in order to increase the tool’s sensitivity. However, not all true positive patients will have risks that can be mitigated by the preventive intervention being offered. In other words, some of the patients correctly identified by the tool as being high risk may still experience an unplanned hospital
admission despite the best efforts of the preventive intervention. In order to reduce waste and improve the efficiency of the preventive programme, it would therefore be helpful to predict the subgroup of the true positive patients who will benefit from different preventive interventions and to restrict each preventive intervention to these “high-risk, high-impact” individuals. A number of tools, known as impactibility models, have been described whose aim is to identify these subgroups. See Table 1.

### Table 1: Types of Impactibility Model

<table>
<thead>
<tr>
<th>Approach</th>
<th>Details</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gap analysis</td>
<td>Using this approach, a CCG would focus attention on those high-risk patients whose care appears suboptimal, such as patients with multiple “gaps in care.” An example of a gap is a patient with heart failure who had not been offered beta-blocker medication despite having no contraindications.</td>
<td>This approach may help reduce health care inequalities because suboptimal care tends to be more prevalent in more deprived areas (the so-called “Inverse Care Law”).</td>
</tr>
<tr>
<td>Focus on impactful conditions</td>
<td>Here, a CCG would prioritise high-risk patients who had a disease or conditions known to be responsive to preventive care, such as patients with an ambulatory care–sensitive condition, such as heart failure.</td>
<td>This approach may help reduce health care inequalities because ambulatory care-sensitive conditions are more prevalent in more deprived areas.</td>
</tr>
<tr>
<td>Exclude highest risk individuals</td>
<td>Some U.S. health care organisations report that they de-prioritise those patients whom they expected to respond poorly to preventive care, such as people with dementia, mental illness, or language barriers.</td>
<td>This approach raises serious ethical concerns, would worsen health care inequalities, and may well be unlawful in the UK.</td>
</tr>
<tr>
<td>Exclude individuals who are unlikely to respond</td>
<td>Some programmes exclude all of the very highest-risk patients because they regard such patients as being less amenable than others to preventive care.</td>
<td>Since very high risk patients have such a high propensity for Triple Fail events, it is generally worthwhile expending resources on few who can be affected is usually worth the effort.</td>
</tr>
</tbody>
</table>
5.2 As we can see from Table 1, CCGs need to be very cautious in their use of impactibility models. All four of the described approaches have the potential to improve the efficiency of the preventive care being offered and the first two types of impactibility models should help reduce health care inequalities. However, other types of impactibility model may worsen health inequalities and must be avoided. While there is no suggestion that any NHS organisations have systematically excluded patients on the basis of factors such as mental illness or language impairment, clinicians do frequently “filter” the high-risk patients identified by a risk stratification tool in more informal ways. For example, clinicians may review the list of high-risk individuals and then select those individual patients that they think are most likely to benefit from the preventive care or patient education, with the remaining patients receiving standard care. CCGs therefore need to ensure that the heuristics used by their well-meaning clinicians are not inadvertently worsening health inequalities, for example by excluding patients with cognitive impairment, alcohol problems, or patients whose first language is not English.

6. Evidence base for preventive programmes

6.1 A review published by the King’s Fund in 2010 found that there was little evidence to support many of the hospital avoidance interventions being offered at the time to high-risk patients. Indeed, a more recent study suggested that many interventions actually increased rather than decreased admission rates.

6.2 However, a 2013 review found good evidence that the following interventions can prevent hospital admissions: continuity of care with a GP, early senior review in the emergency department, structured discharge planning, advanced care planning, and coordination of care at the end of life. There is also evidence that highly structured programmes, such as the GRACE programme, can improve the quality of care and reduce acute care utilisation. Furthermore, for frail older people who are currently in hospital, there is good evidence that multi-dimensional geriatric assessment improves quality of life while lowering mortality, readmission rates, long-term care use and total costs.

6.3 Overall, however, there is little robust evidence that many of the programmes currently being offered to high-risk strata improve outcomes while reducing costs. There is therefore a pressing need for further research and evaluation.
The Way Forward

7. Safeguards for NHS organisations

7.1 As we have seen, risk stratification tools offer a potential means of addressing some of the most pressing challenges faced by the NHS. With this great promise, however, comes a range of potential risks and therefore a series of safeguards is needed, both ethical and scientific.

7.2 The first step for an NHS organisation interested in beginning or expanding its risk stratification programme is to conduct an opportunity analysis. This process involves analysing population data to identify the incidence of low-quality, high-cost, poor-experience events such as unplanned hospital admissions, readmissions, receiving overly invasive treatment for a preference-sensitive condition, and receiving over-medicalised care at the end of life.5

7.3 The next step is to determine the ethics of predicting these adverse events and offering an intervention designed to prevent them. As we have seen, any risk stratification programme has the potential to cause more harm than good. The World Health Organisation published ten prerequisites that should be met by any ethical screening program.24 Because risk stratification is analogous to population screening, it has been argued that equivalent caveats should apply; therefore, these criteria can form a useful basis for an ethical review.5 See Box 3.

Box 3: Prerequisites for Risk Stratification

1. The event being predicted should be an important health problem.
2. There should be an accepted intervention offered to high-risk patients.
3. Resources and systems should be available for timely risk stratification.
4. There should be sufficient time for intervention between risk stratification and the occurrence of the adverse event.
5. A sufficiently accurate predictive risk model for the event should be available.
6. The risk stratification tool should be acceptable to the population at large.
7. There should be an accepted policy about who should be offered the preventive intervention.
8. The natural history of the adverse event should be adequately understood by the organisation offering the preventive.
9. The cost of risk stratification should be “economically balanced,” (i.e., it should not be excessive relative to the cost of the program as a whole).
10. Risk stratification should be a continuous process, not just a “once and for all” occurrence.
7.3.1 As part of this ethical review, an NHS organisation would need to consider the information governance implications, the predictive accuracy of the risk stratification tool, and the effectiveness of the preventive intervention that is offered to high-risk patients.

7.4 As we have seen, the literature on cost-effective preventive interventions is patchy. Accordingly, it is essential that any preventive programmes be evaluated properly – either as part of a formal research study or through local service evaluation and clinical audit. With any evaluation, it is important to establish a valid comparator group. Because of the phenomenon of \textit{regression to the mean}, a pre-post study does not constitute a valid comparator group.\textsuperscript{14} Instead, CCGs should consider using techniques such as \textit{pragmatic randomised controlled trials},\textsuperscript{25} \textit{propensity score matched cohort studies}\textsuperscript{20} or \textit{regression discontinuity analyses}\.\textsuperscript{26} Service utilisation and cost savings are likely to be key outcomes of interest for any evaluation, but other factors – such as patient experience, health outcomes and health inequalities – should also feature.

7.5 In addition, it is important to use the data generated in any risk stratification programme as a form of continuous feedback loop to improve the performance of the programme. For example, a regression analysis might show that patients with certain characteristics were more likely than others to benefit from the preventive intervention being offered. This insight should be used to build or adjust an impactibility model to ensure that patients with these characteristics were prioritised, unless that adjustment violated ethical considerations.\textsuperscript{5}

7.6 Finally, one of the problems that can hamper both evaluations and feedback loops is the issue of small numbers. For example, an evaluation study typically requires data for several hundred patients in order to detect any significant differences. For this reason, NHS organisations should consider working collaboratively with each other by implementing a common risk stratification tool and a standardised preventive intervention then pooling their data for analysis. At a local level, Academic Health Sciences Networks may have an important role to play here; and at a national level, both the Health Foundation and the Nuffield Trust have established programmes of evaluation that involve networks of NHS organisations from across the country.

7.7 At an international level, the European Commission’s \textit{Activation of Stratification Strategies and Results of the Interventions on Frail Patients of Healthcare Services} project (\url{www.assehs.eu}) is bringing together risk stratification professionals from health services, academia and research centres from across the European Union to study current existing health risk stratification strategies and tools and the challenges to spread their use and the application on frail older patients.
Conclusion

As we have seen, risk stratification is a topic of intense research and it is clear that this is not, by any means, the last word on the subject. There is so much more to learn that this topic will remain a live one and NHS organisations will need to continue to engage with.
References


