

Predictive risk and health care: an overview

Research summary

March 2011

About the authors

The Research Group at the Nuffield Trust have considerable experience in the application of predictive risk modelling techniques in the UK. This includes work on the development of case finding tools such as PARR (Patients At Risk of Re-hospitalisation) and the Combined Model; a feasibility study of models that predict future use of social care; work on person-based resource allocation; and a range of national evaluations of interventions to reduce hospital use, including the Whole System Demonstrator Project, integrated care pilots and selected Partnership for Older People Projects (POPP) and Virtual Wards.

The team includes:

Theo Georghiou: Senior Analyst

Ian Blunt: Senior Analyst

Adam Steventon: Senior Analyst

Geraint Lewis: Senior Fellow

John Billings: Visiting Fellow

Martin Bardsley: Head of Research

Risk adjustment has been used for many decades in research studies, but since the 1980s the approach has been increasingly used in managing health systems, and looks set to play a greater role in the NHS in the future. In this research summary, we review how predictive risk adjustment techniques are currently being used in the NHS, explore some of the challenges involved in applying these techniques in practice, and examine some recent advances and new developments in this field.

Key points

- Risk adjustment tools are used widely in the US and in mainland Europe to help determine health payments; either for fixing ‘capitated’ budgets or for deciding reimbursement rates for individual patients.
- To date, the most widespread use of risk adjustment in the NHS has been in the use of ‘case finding’ predictive modelling tools, such as PARR (Patients At Risk of Re-hospitalisation) and the ‘Combined Model’. Primary care trusts (PCTs) and GP commissioning consortia are expected to make increasing use of such tools to stratify the health risk of the populations they serve.
- Current work on a person-based resource allocation (PBRA) formula for the NHS in England will see risk adjustment being used to set GP practice and possibly GP consortium budgets, based on the predicted per capita inpatient and outpatient costs of their registered populations.
- Researchers at the Nuffield Trust are also using the principles of risk adjustment to evaluate complex community-based interventions to reduce avoidable hospitalisation, for example by testing the cost-effectiveness of telehealth and telecare devices, ‘Virtual Wards’ and national integrated care pilots.
- The increasing ability to link large datasets at an individual level pseudonymously means that the range of data used on these types of models, and their applications look set to grow. For example, recent work has demonstrated that it is possible to build predictive models based on linked GP, hospital and social care information to predict future costs of social care.



Find out more online at: www.nuffieldtrust.org.uk

Background

The use of predictive risk models in health care was developed as a technique to take account of the fact that different people use markedly different amounts of health care services.^{1,2} The distribution of health care use across a population tends to be highly skewed, with a small number of people accounting for a large share of total resources.³ This distribution has important implications for how preventive services are targeted, how budgets are set, and how health services are regulated and evaluated.

Predictive risk adjustment tools use relationships in historic, routinely-collected electronic health data to determine the expected future health care resource use of each individual person in a population. The process of building a risk adjustment tool typically involves analysing millions of individual health records. Patients who are similar with respect to a predicted future event, such as admission to hospital, are then grouped together. Once a risk adjustment tool has been built, it can be applied to a similar population to estimate expected future events for people at different levels of risk. The information it generates can allow better comparisons to be made between health care providers, and provides commissioners with more accurate estimates of likely future costs.⁴

'Dumping' and 'cream skimming'

Researchers have used the principles of risk adjustment for many decades, for example when epidemiologists adjust for the age and sex mix of a population. In health care, the routine use of risk adjustment changed in the 1980s and it now represents a multi-million dollar industry in the US. This expansion in the use of the technique was triggered by the observation that Medicare health maintenance organisations (HMOs) in the US were, on average, enrolling healthier, lower-cost individuals than the average Medicare recipient in order to make a surplus and offer competitive premiums for care.⁵ Unlike traditional fee-for-service Medicare, HMOs received an annual prospective payment that was used to pay for all of the health care needs of their enrolled population. This meant that HMOs had an incentive to attract low-cost individuals in order to keep their expenditure low. The observation that HMOs were attracting, or 'cream skimming', low-cost patients meant that more sophisticated methods were needed to ensure that health care funding remained equitable. Such methods would also help deal with the issue of 'dumping', whereby an organisation seeks to exclude enrolled patients who might be expected to have higher than average health care costs.

In the UK there has been no competition between commissioning organisations (like HMOs) for patients, so there was less impetus for developing risk adjustment techniques in the UK. A 2001 review by Majeed and colleagues noted how far the UK lagged behind the US in this field;² in fact, the first major applications of risk adjustment in the NHS were not used in the funding of commissioning organisations, but rather for 'case finding' individual patients at high risk of unplanned hospital admission. In this context, risk adjustment is used to identify which patients are expected to incur high costs in the future, so that these individuals can be offered preventive care 'upstream' in the hope of making net savings 'downstream'. More recently, however, risk adjustment techniques have been applied to the funding of NHS commissioning organisations in England through the PBRA formula.¹⁹ Risk adjustment is now also increasingly being used to evaluate preventive services aimed at averting unplanned hospital admissions.

Accessing information for modelling

In order to develop predictive models, it is necessary to link individual health records over time and often across different datasets. To protect patient confidentiality, this process is generally conducted pseudonymously (see Box 1). The ability to construct individual patient histories that span different health and social care databases adds a new dimension to the analysis of routine data allowing analysis of cohorts over time.⁶

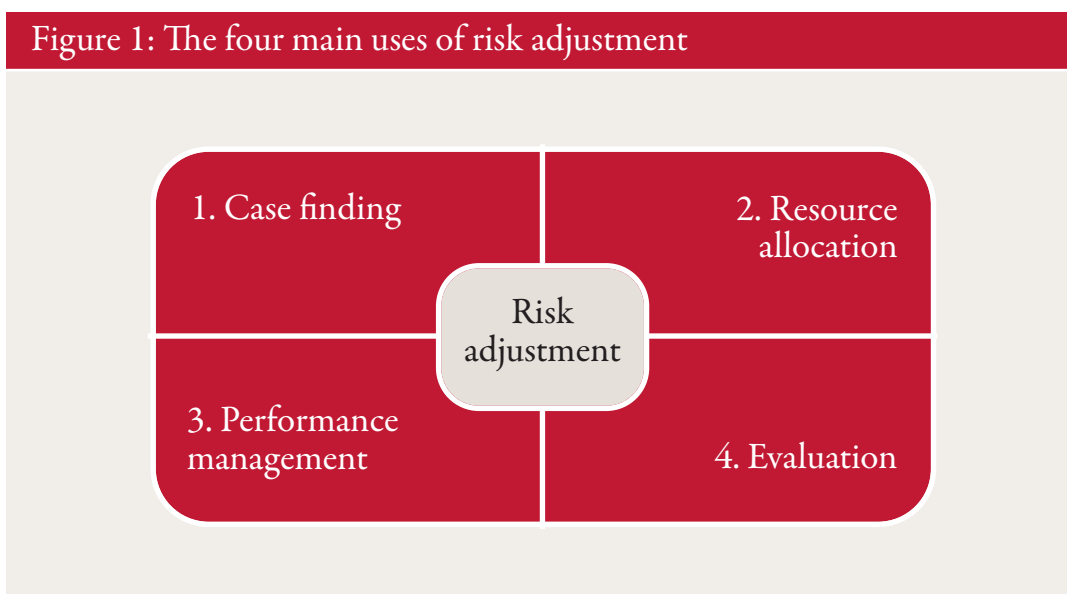
Box 1: Protecting confidential information

The UK Government has made it clear that a fundamental principle governing the use of person-identifiable information by any part of the NHS or the research community is that of 'informed consent'. However, the size of the datasets required for predictive modelling generally means that it is not feasible to seek individual consent from all members of a population to use their data for modelling. Normally, in situations where consent is unobtainable, no information that identifies individual patients may be used.

The Patient Information Advisory Group (PIAG) issued a ruling in 2006 that the principles of confidentiality could be met by encrypting data in such a way that they became effectively de-identified. This is achieved by rendering the data pseudonymous by contracting certain fields (for example replacing date of birth with year of birth), removing sensitive fields (such as the name and address) and encrypting the NHS number with a suitably secure algorithm so that it becomes a meaningless pseudonym.

There is no single standard approach to the process of risk adjustment, but different approaches include statistical techniques such as multiple regression and neural networks. In the US, there is a wide range of competing products ('risk adjusters') available, both commercial and non-commercial, and the Society of Actuaries regularly compares the accuracy of different risk adjustment tools based on their performance on a test dataset.⁷ These tools are now finding application within the UK, for example, several PCTs are using products such as the ACGTM system, developed at Johns Hopkins University, or the DxCGTM classification developed by Verisk, and used widely in Europe.

Applications of risk adjustment

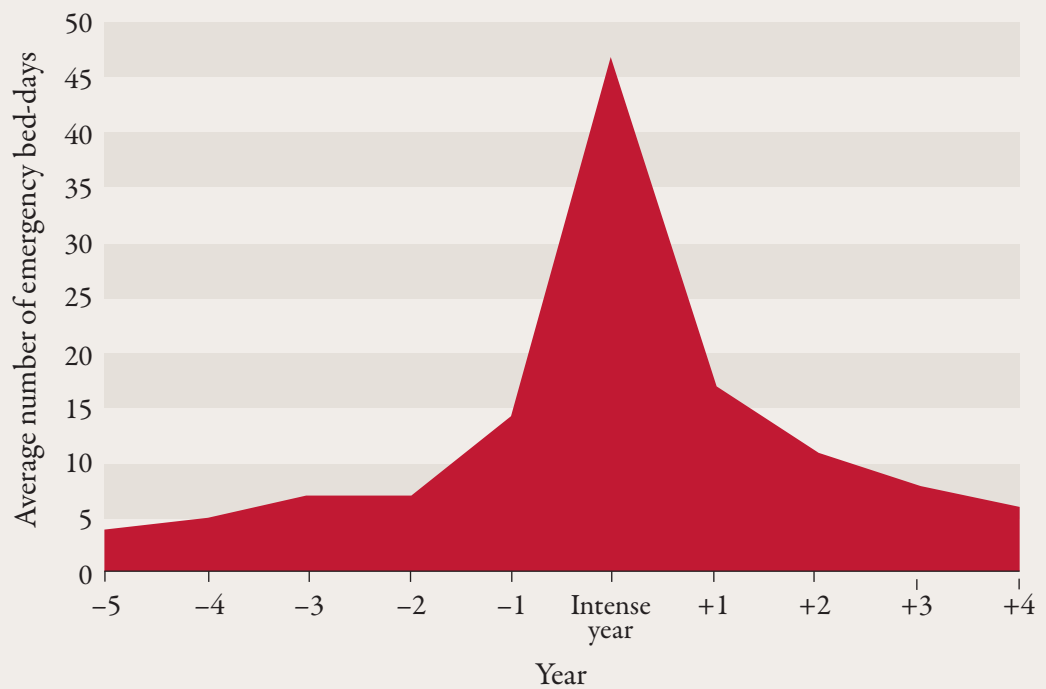


1. Case finding

Health care costs are highly skewed across a population, so high-cost patients could in principle be offered intensive, costly 'upstream' preventive care to improve their quality of life, that might yield significant net savings for the health service from averted 'downstream' costs.⁸ However, offering preventive care to patients who are currently experiencing multiple hospital admissions can be inefficient because, even without intervention, such patients will on average have fewer unplanned hospital admissions in the future. This phenomenon is called 'regression to the mean' (see Figure 2), and it implies that hospital-avoidance interventions are best offered according to the future, not current, risk of hospitalisation of an individual. In the UK there is growing experience of using predictive models for case finding purposes. An example of a tool to identify patients at risk of future re-hospitalisation is the PARR model.⁹ This tool, which was commissioned by the Department of Health, uses existing inpatient datasets to ascribe a risk 'score' to patients that reflects their chances of experiencing a hospital readmission in the coming 12 months. PARR is now widely used across England to identify patients who should be offered hospital-avoidance support, such as care from a community matron or admission to a virtual ward. Another case finding tool is the 'Combined Model', which combines data from GP practice records ('Read code' data) and hospital data to predict emergency admissions (see Box 2).¹⁰

A recent extension of this approach at the Nuffield Trust has shown that it is possible to link health and social care data at the individual level in order to predict social care costs.¹¹

Figure 2: Regression to the mean



Source: Department of Health analysis of Hospital Episode Statistics for England

Figure 2 shows a ten-year extract of Hospital Episode Statistics for England. A cohort of frequent hospital users was identified in the intense year (i.e. year of increased hospital use), and the hospital usage of this cohort of patients was tracked for five years beforehand and five years afterwards. The rapid reduction in bed-day usage seen in year +1, which occurs without specific intervention, illustrates the phenomenon of regression to the mean and demonstrates the importance of identifying people early.

Box 2: Case finding tools based on predictive risk models developed for the NHS

Patients At Risk of Re-hospitalisation

The PARR model was built using a ten per cent sample of Hospital Episode Statistics (HES) data together with variables from the national census. The model can be downloaded and run free of charge by any NHS organisation.⁹ A survey conducted in 2007 suggested that over 70 per cent of the 96 PCTs that replied to the survey were using PARR.¹² Originally, PARR came in two versions: PARR1 made predictions about ambulatory care sensitive admissions; whereas PARR2 made predictions about all admissions regardless of diagnosis.¹³ A new version of PARR, called PARR+, has removed this distinction.

Combined Model

A disadvantage of PARR is that it is unable to make any predictions about the vast majority of the population who have not had a recent hospital admission. The development of a Combined Model represented the final stage of the predictive modelling project funded by the Department of Health. The Combined Model is so named because it was designed to run on a combination of hospital data (from the Secondary Uses Service) and GP (Read code) data. The Combined Model is able to generate a risk score for every member of a registered population that reflects their risk of having an unplanned hospital admission in the next 12 months.

PRISM

In 2010, the Welsh Assembly Government launched the Predictive Risk Stratification Model (PRISM). This tool, which was developed by Health Dialog, is analogous to the Combined Model in that it is run on a combination of hospital data and GP data. An important difference, however, is that PRISM is hosted centrally by the NHS Wales informatics services and results are made available to GPs through a secure website called the Welsh Predictive Risk Service.¹⁴

SPARRA

In Scotland, the Information Services Division (ISD) of NHS Scotland developed a predictive model that is analogous to PARR. Known as the Scottish Patients At Risk of Readmission and Admission (SPARRA), the model is currently run centrally by ISD and results are sent on a quarterly basis to each Health Board.¹⁵ ISD have also developed a model dedicated to predicting admissions to psychiatric hospitals, known as SPARRA-MD.¹⁶

2. Resource allocation

Health care systems that use prospective payment systems such as capitation are susceptible to the phenomenon of ‘cream skimming’, whereby providers or health plans preferentially pick the most profitable patients.¹⁷ Skimming (and its opposite, ‘dumping’, see page 4) undermine the efficiency and quality of care. As a consequence of these perverse incentives, the assessment of ‘risk’ in the US has become a routine process for health insurers and providers alike.

Similar risk adjustment approaches can be found in the funding mechanisms for health systems in certain European countries such as The Netherlands and Germany.¹⁸ More recently, the Department of Health in England has funded the development of a person-based, risk-adjusted resource allocation system for the NHS. In this project, pseudonymous health care records were linked and used to categorise individual patients into different groups based on their expected health care needs. These groups, together with a range of additional variables, can now be used to predict future health costs at an individual level. The aggregated predictions for a GP practice list are now being used to set GP practice budgets for commissioning so that they reflect the underlying morbidity of the patients registered at that practice.¹⁹ This approach, known as PBRA, uses a model that was built using three years’ worth of historic, pseudonymous hospital data. Variables in the first two years, including the diagnoses of patients registered at each practice, were used to predict the costs of providing care in the third year. The results were used to compute practice-specific weightings for a range of age and sex categories. These weightings can now be applied to the most recent population data in order to determine a notional budget for each GP practice, reflecting the expected expenditure for its registered patients in the coming year.²⁰ One of the advantages of this approach is that it does not assume that all patients live in the same geographic area, meaning that it is well suited to estimating notional budgets for GP practices and potentially future commissioning consortia. Further work is underway to refine these models and the ways in which they are applied in practice.

3. Performance management

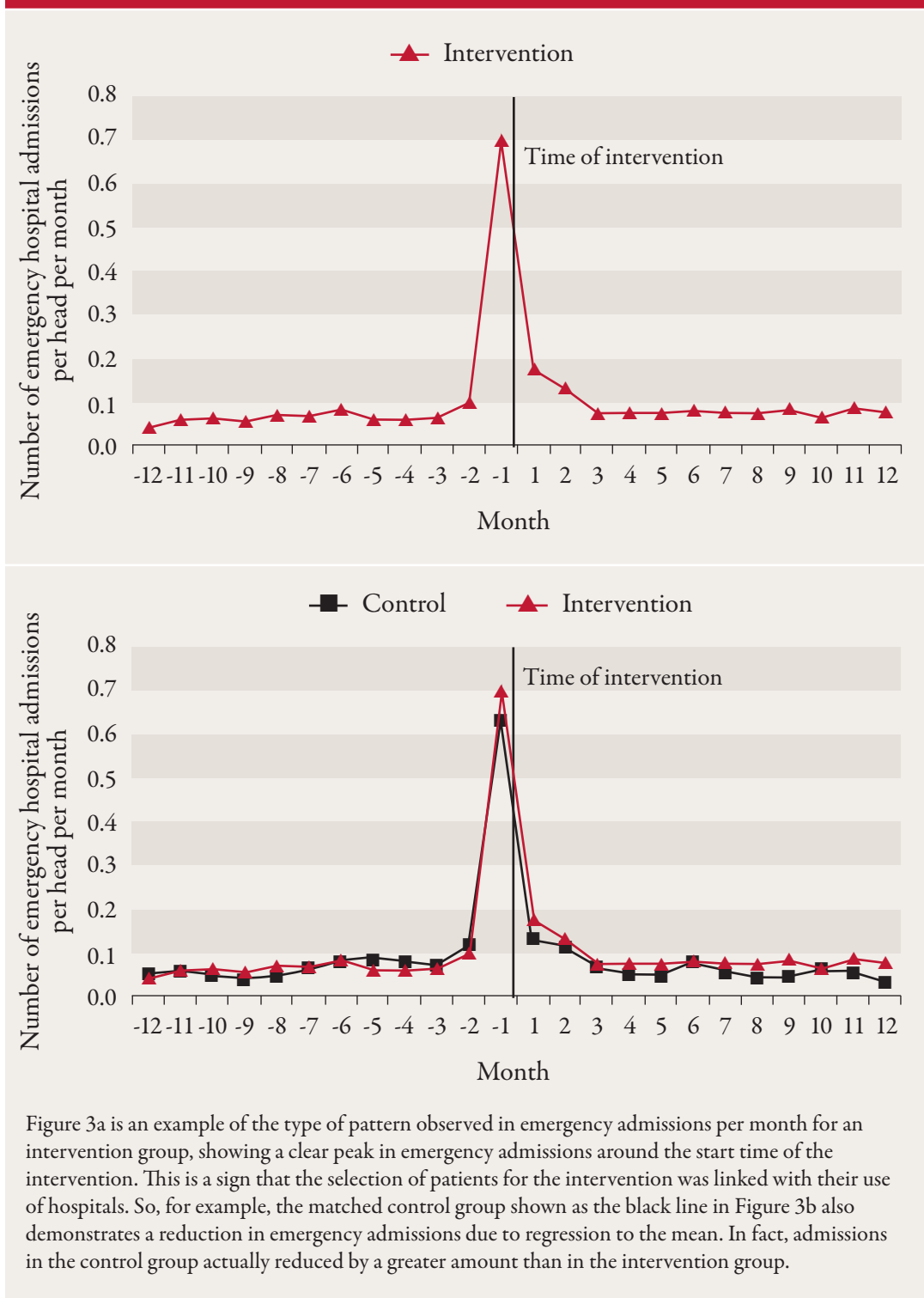
In order to make fair comparisons, managers and regulators need performance indicators that are standardised to ensure that any observed differences in performance are not due to factors beyond the control of the organisation under scrutiny. Without risk adjustment, the danger is that an apparent reduction in, say, emergency bed-days may arise not because of better health care provision in the community, but because of a change in the case mix of patients. By risk-adjusting for the characteristics of the population under consideration, regulators can make more meaningful comparisons of health care providers and commissioners. This is particularly true for ambulatory care sensitive conditions, which are a group of diagnoses that should not result in hospitalisation if managed well in the community.²¹

4. Evaluation

The final major application of risk adjustment techniques is in evaluating interventions, for example those aimed at maintaining the independence of older people with chronic illnesses. Investment in such services is often predicated on the impact that they can have on reducing hospitalisation rates and the use of social care. Yet measuring the impact of interventions on service use can be challenging, particularly if they are being targeted at people who have had recent encounters with the health care system. In these cases, a reduction in hospital utilisation may occur simply as a result of regression to the mean, so evaluations that do not take account of this phenomenon may be misleading.²² By analysing linked datasets, however, it is possible to track the health and social care use of individual pseudonymous patients before and after an intervention. The observed use of health and social care services can then be compared with expected rates as determined by different risk categories.

Researchers at the Nuffield Trust are now using this technique to evaluate the impact of a range of interventions on health and social care. Interventions include the Whole System Demonstrator pilot of telehealth and telecare, the national integrated care pilots, and retrospective evaluations of selected POPP initiatives and Virtual Wards.

Figure 3a and 3b: Example of analysis of control group and intervention group



Limitations of predictive models

Building a predictive model requires very large amounts of routine data. Certain datasets, such as HES in England, are relatively straightforward to obtain, and have already been cleaned ready for analysis. Obtaining more detailed local datasets can be more logistically burdensome and relies on the help and goodwill of local analysts. The accuracy of predictive models is inevitably constrained by the range and accuracy of information contained in computerised datasets, and therefore their accuracy is always far from perfect. Moreover, when a predictive model is being designed for setting budgets, or for determining remuneration, analysts must exclude certain variables in order to avoid perverse incentives. For example, hospitals that knew that their budget next year could be influenced by the recorded illness severity of their patients, or the number of procedures performed, might be tempted to 'up-code' their patients or perform more tests than necessary. Analysts therefore exclude these types of variable; however, this reduces the predictive power of the models still further.

Emerging developments

As well as the NHS predictive models discussed in Box 2, a range of proprietary models are now available for purchase by NHS organisations, including several models that were developed originally in the US. Other developments include:

Social care

There is high-quality evidence that interventions such as domiciliary multi-dimensional assessment with frequent follow-up can successfully prevent or delay admissions to care homes.^{23,24} In 2008, the Department of Health commissioned the Nuffield Trust to examine the feasibility of constructing such a model.¹¹ Modelling with social care data presents a number of challenges. For example, the conventions and classifications in place for social care data are much less standardised than those for health care data; social care data do not typically include the NHS number; and there are different information governance rules for social care. Nevertheless, there is a growing interest and some experience in exploiting linked health and social care data to predict future social care use,⁶ or to describe care pathways spanning the entire health and social care economy.²⁵

Predicting impact

One problem with predictive models is that some of the high-risk patients identified may not, in fact, be amenable to 'upstream' preventive care. In the US, insurers and disease management organisations are eager to invest resources only in those patients who are likely to benefit. For this reason, 'predictive impactability models' are being developed to identify the subset of at-risk patients in whom preventive care is most likely to succeed.²⁶ One way to improve impact may be to prioritise patients with ambulatory care sensitive conditions. Another approach is to prioritise patients with multiple 'gaps' in the quality of care (a 'gap' being a disparity between evidence-based guidelines and the care recorded in routine data).²⁷ Both of these approaches may be expected to help reduce health care inequalities, since higher prevalence of ambulatory care sensitive conditions and lower quality care can be associated with more deprived populations and areas. Other applications of impactability modelling, however, may worsen inequalities if they are allowed to develop unchecked. Examples are impactability models that exclude patients with mental health problems, or those that exclude patients with poor English language skills.

Improvements in information

The process of collating linked pseudonymous datasets for predictive modelling can lead to a number of useful by-products, as the models require information to be structured along an individual's pathway of care. Information aggregated in this way can be used to create useful 'dashboards' for use by commissioning organisations to monitor the quality, access and cost-effectiveness of care. However, to have the maximum impact at a clinical level, these types of data will need to be 'pushed' into the electronic medical record rather than requiring doctors to 'pull' the information from separate sources. We suspect that the integration of risk scores within clinical information systems is likely to improve in the future.

Shorter-term predictors of readmissions

Tools such as PARR and the Combined Model make predictions over a 12-month prediction window, that is they identify patients at risk of admission or readmission in the coming year. More recently, researchers in Canada have developed a predictive model with a much shorter prediction window. The LACE model is designed to be applied to current inpatients on their day of discharge from hospital, and predicts their likelihood of readmission within 28 days.²⁸ Such predictions could be especially useful to NHS hospitals given current plans whereby hospitals may not be reimbursed for certain readmissions within 30 days. Analysts at the Nuffield Trust are currently validating the LACE tool for NHS data and optimising the model's performance for use in England.

Future agenda

Some of the key developments emerging in relation to risk adjustment in the UK include:

- Developing a better understanding of how different risk stratification systems perform on NHS data and their potential applications, strengths and weaknesses. The approach taken by the Society of Actuaries in the US might be helpful here, i.e. an independent organisation comparing the performance of several competing risk classifications on a set of test data.
- Clarifying the procedures for handling linked datasets in ways that do not breach patient confidentiality. This is especially important for case finding applications, where an individual's treatment may be affected by their risk score.
- Developing information systems that are accurate, up-to-date and capable of combining the relevant information from a range of health and social care sources. In England, the uptake of the Combined Model has been restricted by problems in accessing computerised patient-level information from GP practices.
- Incorporation of new datasets (for example, social care data, housing data, community health services data) that may contain additional predictor variables.
- Development of impactability models, especially those that prioritise patients with quality 'gaps' in their care.

References

1. AcademyHealth, *2008 HSR Impact Awardee: Improving the Financing and Delivery of Health Care with Risk-Based Predictive Modeling* (2008).
www.academyhealth.org/files/awards/Risk-BasedPredictiveModeling.pdf
2. Majeed A, Bindman AB and Weiner JP, 'Use of risk adjustment in setting budgets and measuring performance in primary care I: how it works' *BMJ* 2001;323(7313):604–607.
3. Cummings NA, Cummings JL and Johnson JN (eds), *Behavioral Health in Primary Care: A Guide for Clinical Integration* (1997) Madison, CT: Psychosocial Press.
4. Cucciare MA and O'Donohue W, 'Predicting future healthcare costs: how well does risk-adjustment work?' *Journal of Health Organization and Management* 2006;20(2):150–162.
5. Biles B, Dallek G and Nicholas LH, 'Medicare advantage: déjà vu all over again?' *Health Affairs* 2004;Suppl Web Exclusives:W4-586–97.
6. Bardsley M and others, 'Predicting who will use intensive social care: case finding tools based on linked health and social care data' *Age and Ageing* 2011;40(2)265–270.
7. Winkelman R and Mehmud S, *A Comparative Analysis of Claims-Based Tools for Health Risk Assessment* (2007) Society of Actuaries.
www.soa.org/files/pdf/risk-assessmenttc.pdf
8. Billings J and Mijanovich T, 'Narrow Model: The Authors Respond' *Health Affairs* 2008;27(3):900.
9. Billings J and others, 'Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients' *BMJ* 2006;333(7563):327.
10. Wennberg D and others, *Combined Predictive Model: final report and technical documentation* (2006).
www.kingsfund.org.uk/document.rm?id=6745
11. Bardsley M and others, *Predicting Social Care Costs: A feasibility study* (2011) London: Nuffield Trust.
12. Personal communication from Natasha Curry, The King's Fund.
13. Billings J and others, *Case Finding Algorithms for Patients at Risk of Re-Hospitalization, PARR 1 and PARR 2* (2006) London: The King's Fund.
14. www.wales.nhs.uk/nwis/news/17929
15. SPARRA: Scottish Patients At Risk of Readmission and Admission. Edinburgh: NHS National Services Scotland (2006).
www.isdscotland.org/isd/files/SPARRA_Report.pdf
16. SPARRA Mental Disorder: Scottish Patients at Risk of Readmission and Admission (to psychiatric hospitals or units). NHS Scotland Information Services Division (2009).
http://showcc.nhsscotland.com/shsv53_STAGE/servlet/FileBuffer?namedFile=SPARRA-MD-report-final.pdf&pContentDispositionType=inline
17. Ellis R, 'Creaming, skimping and dumping: provider competition on the intensive and extensive margins' *Journal of Health Economics* 1998;17:537–555.
18. van de Ven WPM and others, 'Risk adjustment and risk selection in Europe: 6 years later' *Health Policy* 2007;83(2):162–179.
19. Dixon J and others, *Developing a person-based resource allocation formula for allocations to general practices in England* (2010) London: Nuffield Trust.
20. Department of Health, *Practice-Based Commissioning Budget Guidance for 2010/11* (2010) London: Department of Health.
www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_111057

21. Billings J and others, 'Impact of socioeconomic status on hospital use in New York City' *Health Affairs* 1993;12(1):162–173.
22. Roland M and others, 'Follow up of people aged 65 and over with a history of emergency admissions: analysis of routine admissions data' *BMJ* 2005;330(7486):289–92.
23. Lewis GH, *Predicting Who Will Need Costly Care: How Best to Target Preventive Health, Housing and Social Programmes* (2007) London: The King's Fund.
24. Stuck AE and others, 'Home visits to prevent nursing home admission and functional decline in elderly people: systematic review and meta-regression analysis' *JAMA* 2002;287:1022–8.
25. Bardsley M, Georghiou T and Dixon J, *Social Care and Hospital Use at the End of Life* (2010) London: Nuffield Trust.
26. Lewis GH, "Impactability Models": identifying the subgroup of high-risk patients most amenable to hospital-avoidance programs' *Milbank Quarterly* 2010;88(2):240–255.
27. Milliman Inc, *Milliman Care Guidelines* (2009).
www.milliman.com/expertise/healthcare/products-tools/milliman-care-guidelines/
28. van Walraven C and others, 'Derivation and validation of an index to predict early death or unplanned readmission after discharge from hospital to the community' *CMAJ* 2010;182(6):551–557.

For more information about the Nuffield Trust,
including details of our latest research and analysis,
please visit www.nuffieldtrust.org.uk

 Download further copies of this research summary
from www.nuffieldtrust.org.uk/publications

 Subscribe to our newsletter:
www.nuffieldtrust.org.uk/newsletter

 Follow us on Twitter: [Twitter.com/NuffieldTrust](https://twitter.com/NuffieldTrust)

Nuffield Trust is an authoritative
and independent source of
evidence-based research and
policy analysis to improve
health care in the UK

59 New Cavendish Street
London W1G 7LP
Telephone: 020 7631 8450
Facsimile: 020 7631 8451
Email: info@nuffieldtrust.org.uk

 www.nuffieldtrust.org.uk

Published by the Nuffield Trust.
© Nuffield Trust 2011. Not to be reproduced
without permission.