The Cumberland Initiative
TRANSFORMING HEALTHCARE QUALITY AND COSTS

EMERGENCY SIMULATION

How modelling is resuscitating NHS Urgent & Unscheduled Care

Pioneers from the Cumberland Initiative are working with clinicians and managers in the NHS to tackle the current crisis in Urgent and Unscheduled Care.

This report draws together learning from test-driving different approaches. It shows how modelling provides a safe, inexpensive way to try new methods, uniting the many stakeholders needed to find and implement sustainable solutions.

Acknowledgements

We are indebted to all those who contributed to and attended the launch of Cumberland Initiative at a special forum, entitled 'Remodelling Urgent and Unscheduled Care' on 2 May 2013 at IQ Winnersh, near Reading, where much of this work was presented.

CONTENTS

1. Introduction

'We are helping to resolve the crisis in Urgent and Unscheduled care.'

Our partners in the Cumberland Initiative are already providing vital insights from modelling new options around Urgent and Unscheduled Care, explains **Professor Terry Young**, CI's co-founder.

Tackling the current NHS crisis in Urgent and Unscheduled Care is a delicate process because so many critical services flow through the Emergency Department and so many confounding factors impact, often indirectly, upon it. Thus, changes in the dynamics of the dialogue between GPs and hospitals, or very small increases in long-stay populations, for instance, quickly develop into unacceptable waiting and overcrowding in Emergency Departments and Acute Medical Units.

A sustainable solution requires a new way for the many care sectors and health professionals to collaborate: from primary care, hospitals, social care and community-based care, as well as commissioners and, of course, patients themselves. This is the route to the right response that brings everyone on board. Then they need to 'work the problem' together, testing out options in a safe environment where mistakes can be made without hurting patients or staff or costing a fortune.

The Cumberland Initiative provides precisely this facility. We build networks of clinicians, managers, modellers, academics and industry specialists to develop a holistic approach to healthcare problems. We explore the interconnections of the whole system and we are starting to make a difference.

In this pioneering document, we highlight how these communities that are now working with the Cumberland Initiative have already come together to improve Urgent Care. You will find, in this report, examples of them modelling change across England and Wales.

First, **Russell Emeny**, one of the NHS experts on the crisis, highlights key learning. Solving the problem, he says, relies on reducing avoidable hospitalisation, focussing on home-based solutions and improving patient flows.

Urgent care in hospitals in south Wales is using a highly modern approach - **Dr Julie Vile** is one of the first academic mathematical modellers to be recruited full-time by the NHS (Aneurin Bevan Health Board) and sets out some brilliant insights. Meanwhile, her boss, **Dr Danny Antebi**, Director of Aneurin Bevan Continuous Improvement Centre, details the importance of modelling to his strategic planning. He discusses how the complexity of care required for key vulnerable patients is a vital issue. Finally, a group of academics, led by **Professor Paul Harper** at Cardiff University's School of Mathematics, sets out modelling work they have done to improve ambulance response times and alleviate strain on Urgent Care.

> In times of change, learners inherit the Earth, while the learned find themselves beautifully equipped to deal with a world that no longer exists.
>
> **Eric Hoffer, writer**

'We build networks of clinicians, managers, modellers, academics and industry specialists to develop a holistic approach to healthcare problems.'

Professor Sally Brailsford from University of Southampton explains how one popular modelling approach – Systems Dynamics –has helped managers in Nottingham to understand the inner workings of their health system. The key learning? In this case, that alternatives to admission were more effective in reducing bed occupancy than was discharge management.

Perhaps one of the most insightful contributions comes from **Mr Andrew Fordyce**, a consultant oral and maxillofacial surgeon – and also Clinical Systems Engineer at Torbay Hospital, Devon. Andrew and colleagues have been concerned to get bed levels right in the hospital. His narrative shows how vital clinical leadership is in forging discussion of options and making modelling work to avoid pitfalls – in this case of cutting beds before implementing measures to reduce length of stay.

Finally, there are fascinating contributions from two of the pioneering companies in this field. First, **Peter Lacey**, from Whole Systems Partnership, sketches how modelling has given managers in Lincolnshire the confidence to invest in a GP presence in A&E, to create intermediate care and to close acute beds. Second, **Claire Cordeaux,** from SIMUL8, shows how simulation is being used to develop innovative funding mechanisms supporting integrated health and social care for people with long-term conditions. That's a prize worth struggling for.

Following on from this learning, the Cumberland Initiative, in the summer of 2013, has initiated a series of fresh projects with a number of Trusts across the NHS to support our first Grand Challenge – 'Remodelling Urgent and Unscheduled Care'. We'll be back later in 2013 with feedback on this major and exciting venture.

Professor Terry Young

Co-Founder, Cumberland Initiative
Chair of Healthcare Systems
School of Information Systems, Computing and Mathematics
Brunel University

2. Key steps to stabilise Emergency Care in England

Rebalancing the NHS demands reducing avoidable hospitalisation, focussing on home-based solutions and improving patient flows, explains **Russell Emeny**, a key NHS advisor on emergency care.

We need to identify clear, immediate actions so that we can do as much as possible quickly to stabilise current problems in emergency care.

In my view, our difficulties are caused by the combined effect of long term trends and many smaller stimuli. This has created a fragile system vulnerable to small impacts. As a result, the system has lost equilibrium, is struggling and recovery is slow.

However, we are not helpless. The NHS needs to implement a number of proven tactics to re-stabilise the system in the immediate term. That will provide breathing space to look at the longer term issues.

Proven tactics that could help to re-stabilise the system quickly include: tackling avoidable hospitalisation; focussing on home-based rather than bed-based solutions; improving patient flow along the pathway and particularly through hospitals.

We know the system is stressed: the four-hour arrival to departure performance in Emergency Departments is the lowest in ten years and 12 hour 'trolley-wait' breaches have increased. Time from initial assessment to start of treatment is growing as is the time from start of treatment to decision to admit.

Hospital occupancy, length of stay and waits for admission are all increasing as are delays in ambulance hand-over. Meanwhile, cost improvement programmes are closing beds, but beds are being reopened in escalation.

There are also issues related to NHS 111, social care and continuing health care delays as well as difficulty discharging into community beds.

> *'The system has lost equilibrium, is struggling and recovery is slow. The NHS must turn to tactical solutions to reduce variation and optimise performance as a short term measure to re-stabilise the system.'*

Second – and very important - the NHS is characterised by considerable variations in its performance around the country: there is a four-fold variation in admission rates for people over 65 years old; there are considerable variations in hospital length of stay for patients under different consultants for the same conditions; there is 10 per cent higher mortality at weekends compared with weekdays; Emergency Department (ED) attendances vary according to proximity to the ED. We know that medicine is slow systematically to adopt good practice, even where proven.

Third, there have been big changes in acute care: a 37 per cent increase in emergency admissions over past ten years, of which only 40 per cent is due to changing demography. The rate of hospital intervention is growing much faster than the rate of ageing with much of growth in short stay admissions.

There are various hypotheses for the phenomenon, including: improved medical technology and knowledge allowing more conditions to be managed; that there is reduced threshold for admission; risk aversion is sometimes blamed among (usually junior) doctors, compounded by less experienced junior doctors managing admissions.

Fourth, there is the aggregate impact of small (negative) affects: the development of NHS 111 has dropped down which may be impacting on ED attendance and admissions; changes in out of hours cover; the impact of the Francis Report; funding problems with social, primary and continuing care.

It seems clear that the combined effect of long term trends, financial pressures, medical practice and many relatively minor system vulnerable to small impacts. The system has lost equilibrium, is struggling and recovery is slow. The NHS must turn to tactical solutions to reduce variation and optimise performance as a short term measure to re-stabilise the system. That's why we have to tackle avoidable hospitalisation, focussing on home-based rather than bed-based solutions and improving patient flow along the pathway and particularly through hospitals.

The principles for improving patient flow are clear. A good step is early senior review of patients, soon after arrival and daily senior review for those admitted. The benefits are considerable. In one study, twice weekly consultant ward rounds in two general medical wards were compared with twice daily ward rounds. Average length of stay on study wards fell from 10.4 to 5.3 days with no deterioration in other indicators (re-admissions, mortality, and bed occupancy).

There is considerable potential for improvement. Only 50 per cent of acute medical units have twice daily ward rounds, and 9 per cent have consultants on-take in blocks of greater than 1 day (RCP 2012). There is considerable scope to reduce mortality by adopting RCP guidance.

' Current performance problems arise from multiple factors and constitute a "wicked problem". But we are not helpless. We need to apply known good practice systematically.'

We also need to focus on discharge. Consistently prioritising discharge activities can significantly reduce length of stay in elective or emergency clinical care pathways. In contrast, prioritising discharge activities only when beds are full may have little impact on patient throughput or average length of stay. Increasing beds may increase length of stay with no benefit to patient throughput.

> "Freedom is, first of all, the chance to formulate the available choices, to argue over them and then, the opportunity to choose."
>
> **C Wright-Mills, sociologist**

The principle of speedy, experienced medical assessment should also be applied in primary care. At the moment, too often, there is the pattern of an elderly patient ringing the GP for a home visit early in the day. The GP comes perhaps late morning after completing surgery appointments. Then there might be a long wait for an ambulance – typically much longer than a 999 call. With all the various delays, patients may be getting treatment that is considerably delayed and reach hospital only towards the end of the day, when staff are leaving and they are likely to be kept waiting for longer. It all makes eventual admission more likely and worsens outcomes and likely length of stay.

Can potential admissions be turned around? Most studies suggest that admissions can be avoided in 20-30 per cent of frail people, aged over 75 years.

'Avoiding admissions in this group of older people depended on high quality decision making around the time of admission, either by GPs or hospital doctors. Crucially it also depended on sufficient appropriate capacity in alternative community services (notably intermediate care) so that a person's needs can be met outside hospital, so avoiding 'defaulting' into acute beds as the only solution to problems in the community.' (Mytton et al. British Journal of Healthcare Management 2012 Vol. 18 No 11).

To make a real impact, we need to focus on key groups: the frail elderly at home; terminally ill patients; people in nursing and residential homes; some specific groups with particular chronic conditions such as heart failure.

To sum up, current performance problems arise from multiple factors and constitute a 'wicked problem'. But we are not helpless. We need to apply known good practice systematically. We also need to understand complex trends and the impact of small effects on complex systems in order to achieve sustainable improvement.

Russell Emeny is Director, Urgent and, Emergency Care Intensive Support Team, NHS IMAS

3. Testing options to fix Emergency Care in South Wales

When A&E problems escalated in two local hospitals, Aneurin Bevan Health Board in south Wales recruited **Dr Julie Vile**, a mathematical modeller, to drill into the statistics and to explore different options. Armed with her findings, **Dr Danny Antebi**, Director for ABCi (Aneurin Bevan Continuous Improvement), is developing the Board's innovation strategy.

Understanding the data, modelling some alternative approaches

By Dr Julie Vile

I recently began working for Aneurin Bevan Health Board (ABHB) as a mathematical modeller, as part of an exciting joint initiative between the Health Board and the Operational Research Department at Cardiff University, to support the detailed financial modelling of projects across Aneurin Bevan and explore the potential systems impacts of different intervention strategies.

When I took up my post, I was introduced to two significantly underperforming A&E Departments within Aneurin Bevan (namely Royal Gwent and Nevill Hall Hospitals), and was first asked to provide evidence to support the 'case for change' of processes within the A&E Departments and back system improvement projects.

Before delving straight into the statistical analysis, I spent some time in the Royal Gwent A&E Department to gain a first-hand appreciation of the system. It became clear that the system was operating near crisis point. A fundamental change was needed in the service to respond to the changing profile and increasing demand for A&E assistance.

Some of the most common phrases I heard were: 'The system is in crisis'; 'The acuity of our patients is increasing'; 'We have seen an increase in the 85+ age group'. I came to the conclusion the problem was far from straightforward to solve, but the main challenges facing A&E could be summarised under 3 main categories: too much demand was being exerted on the system; the patient pathways through the department were cumbersome; and there was lack of capacity to take patients out of the system.

' Is a "one size fits all' A&E service really suitable, or should we be thinking about planning a separate service to deal with elderly / frail patients?'

I began by analysing historic demand and activity data, and interviewed clinicians at the department to gain a first-hand appreciation of the system. All of these staff members commented that they had seen a large increase in the number of attendances at A&E over their working career. From analysing activity data from Royal Gwent and Nevill Hall A&E Departments between 2008 and 2012, I confirmed that demand for A&E services had risen – actually by around four per cent over the five year period (equating to around ten extra attendances each month). Yet the overall increase wasn't as dramatic as the exponential growth indicated by clinicians, which motivated me to look more closely at the effect of different age bands.

When analysing demand by age, I found that demand from all the age groups below 50 had, in fact, either remained stable or actually slightly decreased over the last five years; but there was a rise in all groups aged 50+. Whilst the increases for the 75-84 and 85+ groups were small in terms of raw numbers, they accounted for the majority of the overall percentage increase in the demand.

> "Everything that can be counted does not necessarily count; everything that counts cannot necessarily be counted."
>
> **Albert Einstein, scientist**

The picture of an ageing population is familiar to all Health Boards, but the elderly cohort is particularly troublesome for ABHB since they commonly incur excessive lengths of stay. For instance, 2012 data shows that only around 6 per cent of A&E attendances were 85+ in ABHB, but patients presenting at A&E in this category had a 56 per cent chance of being admitted, compared with an average of 21 per cent for other age groups. This is unsurprising since elderly patients commonly have multiple co-morbidities and complex needs. Furthermore, while only 42 per cent of medical patients were 75 years or older, they used 58 per cent of the total number of bed days.

Clinicians mainly attributed this occurrence to delays in social services setting up community care packages. However, patient harm can arise from unnecessarily long stays in hospital, since, when on the ward, already frail patients are prone to lose the capacity to perform everyday tasks such as cooking, washing and cleaning, not to mention the extra possibility of picking up extra illnesses or infections.

Thus, whilst increasing demands in elderly patients presenting at A&E is common to all Health Boards in the UK, ABHB is unusual in terms of the excessive number of bed days taken up by elderly patients. So the problem is not necessarily the raw numbers of patients that are turning up at A&E presenting unscheduled care needs, but the way we are addressing them in our system.

Hence there is an urgent need for revised service configurations to deal with this cohort of patients (such as better community services or long-term care packages etc) to reduce the strain they place on hospitals – especially as the size of this cohort is expected to dramatically increase over the next decade. Predictions indicate that the number of elderly patients with dementia is due to rise exponentially in the next 20 years. This expectation begs the question of whether a 'one size fits all' A&E service is really suitable, or if we should be thinking about planning a separate service to deal with elderly, frail patients.

In addition to evaluating the effect of elderly patients, I have also examined to what extent the demand for A&E services is justified and what proportion of the historic demand could be classed as 'appropriate'.

I found that just over 40 per cent of A&E patients in 2012 were not admitted to hospital or given a follow-up appointment. Maybe not all of these can be classed as 'inappropriate' (since some patients may just have needed assistance to manipulate a dislocated shoulder back into the socket, for example). However, it is likely that a number of these attendances could have been avoided. To encourage fewer patients to present at A&E in the future however, there is a clear need for the community model to become better established and better communicated to patients so they use other services instead of automatically ringing 999 or presenting at A&E.

Another sensible admission reduction strategy could involve revising the role of the Welsh Ambulance Service. Rather than being mainly a transport service, it could be encouraged to treat more patients at the scene of the incident.

Continued >

Moreover, a small pilot study has shown that the introduction of a consultant at the front end of A&E to re-divert demands to more appropriate services could be invaluable. It would be interesting to evaluate the savings that could be achieved from implementing such a practice since the case study (involving placing a consultant at the front end of A&E to answer the phone to GP referrals for one single day) showed that it was possible to reduce the patients brought into A&E via this route by 50 per cent.

The consultant achieved this reduction by diverting some patients to other hospitals, others to community services, strictly telling a handful that they didn't need to come in at all, and deferring others until a quieter slot the next day. The time required to achieve this reduction equated to only 55 minutes on the phone. However, the consultant commented that he still could not afford to do this normally on top of his every day duties, and so would like to see additional capacity set aside for a senior staff member to perform this role.

So there is promising potential in some of the above proposals to reduce demand. However, clinicians emphasised that, no matter what policies implemented to reduce demand in the past, numbers have continued to rise. They see the Health Board as having little influence over emergency arrivals from day to day, or hour to hour.

However, while it may be difficult to gain control over these daily levels, forecasting can provide a good indication of what the demand is likely to be throughout the day/week, as there is a lot of predictability in the arrival pattern of patients.

' In crude terms, one could say that non seven-day working is effectively killing people. This surely justifies changing working patterns'

My recommendations are that we should next concentrate on modelling:

- Impact of an alternative pathway for elderly/frail patients
- Co-locating Minor Injuries Units with a separate entrance point in A&E
- Improving computing facilities by introducing hand held monitors and virtual wards to focus peoples' attention on who is in a bed who shouldn't be there
- Patient boarding

Mathematical modelling could also be used to aid hospitals to manage elective arrivals and discharges. I hope to show that, if consultants began working at 8am daily as they do in CDU, instead of 9am, this could lead to earlier patient discharges and reduce overcrowding in A&E. Private hospitals bring people in as close to their appointment time, and discharge them, as early as possible, so why can't we do the same in the NHS?

To really make an impact, it may also be necessary for consultants to work weekends. We have evidence of a direct relationship between breaches of treatment time targets and the death of more patients on weekends. In crude terms, one could say that non seven-day working is effectively killing people. This surely justifies changing working patterns.

Also, wards could take some responsibility and share the risk when a hospital reaches high levels of escalation (this means that the hospital is operating at a sub-standard performance level and in breach of several standard targets, so additional measures must be taken by staff to try and return to safe operating levels). One strategy which is successful in England, but not yet in operation in ABHB, is for wards simply to take on one extra patient each in a temporary bed or trolley to 'board' them. This typically encourages ward staff to discharge patients as quickly as possible, so treated A&E patients can be moved from the cubicles into the long-stay ward, promoting flow within A&E. As I continue my work for ABHB, I hope soon to model and predict some of the savings that could be achieved from implementing these strategies, alongside service reconfigurations.

Getting the strategy right
By Dr Danny Antebi

The problems are clear - repeated escalation, clinical incidents going up, stories of poor care, queuing ambulances. We need to make a case for change and ensure the organisation has the right focus and a set of ideas that we can all agree on. We cannot make change without bringing on board the local authority, the third sector and the public. Making the case needs the right data, the capacity to win hearts and minds and an assurance of patient safety. Modelling is helping us to make the case.

Leadership is vital. The Board has to be clear about what the agenda is. We can't fix A&E on its own – the health of A&E services is a function of all the systems around it. When flow goes wrong, it affects the whole system. My approach to the board is to emphasise that urgent care and elective care are our core business.

'A service for a 25 year old with broken leg is not going to work well for someone over 85 with co-morbidities and dementia. A lot of problems caring for older people are about complexity science not rocket science.'

> "Apathy can be overcome by enthusiasm, and enthusiasm can only be aroused by two things: first, an ideal, which takes the imagination by storm, and second, a definite intelligible plan for carrying that ideal into practice."
>
> **Arnold Toynbee, social reformer**

Our agenda is focussed on a number of issues. The first is patient flows - poor flow harms and kills patients. It wastes resources, demoralises staff and impacts on other departments.

Second, we need to understand that problems in healthcare can be represented in three ways – as simple, complicated or complex.

1. A simple problem can be dealt with by following a recipe or protocol.
2. A complicated issue is, for example, sending someone to the moon. It has a complicated but clear process. A health example might be acute chest pain. It requires expertise, latest equipment, excellent processes, minimal collaboration and a safe environment. It requires a complicated pathway and organisational ownership.
3. In contrast, an example of complexity would be raising a child – there is no certain way of doing it and no guarantee of success. A health example might be chronic cardiac failure and cognitive impairment. The person needs to stay at home, with responses, support and advice when needed. The patient needs a lot of collaboration, engagement with family, friends and others and occasional high tech interventions. It calls for a complex approach and shared ownership.

This way of looking helps us to understand that a service for a 25 year old with a broken leg is not going to work well for someone over 85 with co-morbidities and dementia. A lot of problems caring for older people are about complexity science not rocket science.

We need to acknowledge that, if we are to prevent people entering hospital, we have to make sure the community is resilient and resourced. Modelling can help us with all of these problems – be they simple, complicated or complex.

Dr Julie Vile is a mathematical modeller, with Aneurin Bevan Health Board, South Wales.

Dr Danny Antebi is Director of Aneurin Bevan Continuous Improvement Centre, ABHB, Wales.

4. Modelling Emergency Medical Services in Wales

A partnership between local health providers and Cardiff University mathematicians and operational researchers is yielding benefits as, together, they tackle ambulance response times and better planning for emergency care, explain **Dr Janet E Williams, Professor Paul Harper** and **Dr Vincent Knight.**

Creating the 'Perfect World Model' for the Emergency Unit

By Dr Janet E. Williams

Cardiff and Vale Health Board in South Wales approached the Operational Research Group within Cardiff School of Mathematics to help determine a number of questions.

First, was further investment in the Emergency Department warranted? Second, could the allocation of resources within the emergency and assessment units be improved? Third, was the capacity within the Emergency Department being used to compensate for problems and constraints that lay beyond the emergency unit?

We mapped out activity in the various part of the emergency unit: the medical and surgical assessment units and A & E. When we analysed the data, we saw that there were three distinct phases in each day as well as three distinct phases in each week in the hospital – Mondays, the rest of the week and the weekend. Majors arrival rates varied across the day and across the week in predictable ways. The GP referral rate also varied, with a particularly high rate on Mondays.

We simulated all the different processes people go through from referral to the time they leave hospital. When you see it running and you can watch people coming in and leaving the system, it is very compelling – it fits people's experience.

' We need champions in the NHS – like Andrew Nelson from Cardiff and Vale, Andrew Fordyce in Torbay and Danny Antebi from Aneurin Bevan Health Board'

With this model, we can look at different scenarios – test out different 'what if' questions. For example, given a certain number of Registrars, Consultants and Nurse Assessors, we can ask how many clinical decision makers are needed. We can see the impact that varying levels of staffing has on the four-hour target. We can identify places where we are achieving targets with fewer staff. We can also take this model and look at beds and levels of occupancy of trolley bays in certain scenarios.

We can show, for example, that targets can be reached even when reducing trolley beds from ten to four. Of course, it must be emphasised this assumes the 'perfect world' – that once the patients are dealt with, they are moved into hospital or discharged. But it helps us to see which hurdles lie in the way of achieving a better healthcare world. If we understand these issues better, we have a better chance to make a difference by pulling the levers that really matter.

We have all this data now. What is needed next? We need champions in the NHS – like Andrew Nelson from Cardiff and Vale, Andrew Fordyce in Torbay and Danny Antebi from Aneurin Bevan Health Board who will take forward what these models can tell us about improvement opportunities.

At the Cardiff School of Mathematics, we are now linked with two of the biggest Trusts in Wales. We hope this partnership will make a real difference.

> "The future is already with us, it is just unevenly distributed."
>
> **William Gibson, science fiction author**

Organising ambulance crews as effectively as possible

By Professor Paul Harper

My own experience of ambulance services first alerted me to the problem. On two separate occasions within the space of a month, we had to wait over 45 minutes before a paramedic arrived to treat my young daughter. It therefore came as no surprise to learn that Wales at that time had the worst ambulance response times in the UK. This led us to begin a project at Cardiff University working with the Welsh Ambulance Service Trust (WAST). Two PhD students have worked with WAST on a range of topics to assist them.

'We found that a forecasting technique called 'Singular Spectrum Analysis' (SSA) was particularly good at producing accurate forecasts.'

It took the service a while to accept mathematicians might be helpful in improving ambulance performance. However, the students worked on forecasting demand, feeding that material into crew rostering of paramedics and understanding where best to locate ambulances to improve response and outcomes.

We initially looked at five years of data across Wales. There had been a five per cent increase in the number of 999 calls to WAST over the five years, but, within that, lots of cyclical patterns. Our challenge was to forecast demand in both the short-term (today, tomorrow, next week) as well as in the medium-term (one month out and beyond). We found that a forecasting technique called 'Singular Spectrum Analysis' (SSA) was particularly good at producing accurate forecasts.

Using SSA, we were able to predict ambulance demand for the next week within between one and two per cent accuracy – and the model works in real time as well. This allowed us to work out demand per shift and subsequently calculate crew requirements per shift. Of course, we had to operate within certain objectives, to minimise labour hours, crew size and overtime, while recognising certain constraints: maximum permitted weekly working hours, maximum night time hours, plus requirements for rest breaks and days off. All of this allowed us to set the optimum shift patterns and rosters for crews, taking into account levels of demand from the forecasts.

Placing ambulances in the best locations

By Dr Vincent Knight

We needed to know the best place to locate ambulances. So we mapped in south Wales where calls come from, along with the location of ambulance stations and hospitals – the three nodes of the network.

Then we built a tool that uses Google to find the travel time matrix – the average time for travel between the nodes, adapted to account for ambulances' capacity to travel faster. The model can be used to locate ambulance so as to meet a hard target such as reaching emergency calls within eight minutes. However, outcome measures are of ultimate importance, so we built into the model the probability of surviving from, for example, cardiac arrest and stroke. That way we could identify the best possible location for emergency ambulances and for rapid response vehicles.

'The service did not necessarily need to buy more ambulances, if it could use its existing stock differently.'

We built a model that can run through a number of scenarios. What if, for example, ambulances were forced to achieve a 20 minute turnaround at emergency units? What if more ambulances are available for each shift?

These different scenarios were very informative. We found, for example, that if ambulances could achieve the 20 minute turnaround (not always possible in some hospitals), then that is as effective - or better - at raising performance than making more ambulances available. It demonstrated that the service did not necessarily need to buy more ambulances, if it could use its existing stock differently.

Paul Harper is Professor of Operational Research within the School of Mathematics, Cardiff University.

Dr Janet E Williams is Senior Lecturer in the Operational Research Group within the School of Mathematics, Cardiff University

Dr Vincent Knight is LANCS lecturer at the School of Mathematics, Cardiff University

5. Finding the best levers for change to improve Emergency and Unscheduled Care in Nottingham.

Modelling healthcare in Nottingham revealed that small changes in the system had big impacts elsewhere and keeping elderly patients out of hospital had the largest effect, says **Professor Sally Brailsford,** from University of Southampton.

In 2001-02, Nottingham City PCT (as was) asked us to help them understand better what was going wrong with emergency and unscheduled care. There was constantly increasing pressure on the system: spiralling demand, rising emergency hospital admissions, cancelled elective operations, long A&E waits - a permanent 'winter crisis'.

A Steering Group had been set up to develop a Local Services Framework for unscheduled care, with membership from all providers: hospitals, ambulance service, in-hours and out of hours primary care, NHS Direct, the Walk-in Centre, social services, community mental health, etc. A team from University of Southampton was commissioned to provide research support, led by Professor Val Lattimer, who is now at the University of East Anglia.

To understand what was going on, we undertook a literature review and compared Nottingham with other health authorities. We also held stakeholder interviews and collected healthcare activity data. We put together a descriptive study of patient pathways and carried out a patient survey focussing on people's preferences.

To model how healthcare was working in Nottingham, we used System Dynamics. This is a powerful way of showing how systems behave, which was developed at MIT in the 1960's by Jay Forrester. The fundamental principle of Systems Dynamics modelling is that the structure of a system determines its behaviour. In other words, the way that the individual components of any system relate to and affect each other determines the overall behaviour over time of the system as a whole. So the system might end up behaving counter-intuitively, in ways that none of the stakeholders actually intend. This is a very important feature for the NHS, where the interaction of complex bits of the healthcare system can have unintended and undesirable consequences – with small factors sometimes having a big impact further down the line.

'What happened if demand at the "front door" (A&E and elective admissions) was altered? How did the system react if emergency admissions were reduced for specific patients, such as elderly people?'

In System Dynamics, we build diagrams showing how sometimes a system gets into virtuous or balanced circles, where the interaction of different elements can lead to stable outcomes. Sometimes, however, it can get into a vicious circle, where the system gets out of balance. The trick is to think of – and then test-drive – changes to the system that can improve behaviour, the system's outcomes and keep it balanced.

One example of a balancing loop involves the relationships between referral rates by GPs to hospitals, bed occupancy rates and waiting lists. If waiting lists rise, then GPs tend to reduce their referrals (dealing with some patients in other ways), so bed occupancy rates fall and so do waiting lists. As GPs gradually see things are getting better, they increase their referrals again, leading to a rise in occupancy rates and waiting lists … and so on. This local part of the whole system has a self-righting balance built into it.

System Dynamics does not model individual patients. However its virtue is that, as a model, it does not allow things to get too complicated. Crucially, it runs very quickly and, by capturing dynamic feedback effects, it will show you the big picture of how the "whole system" interacts – what's the effect on 'B' of doing 'A'. You may be surprised to discover that 'A' also has an effect on 'H'.

The system we designed assumed there would be 4 per cent annual growth in emergency admissions and 3 per cent growth in elective admissions. We then looked at how it worked under various scenarios. What happened if demand at the "front door" (A&E and elective admissions) was altered? How did the system react if emergency admissions were reduced for specific patients, such as elderly people? How did things change if patients were discharged earlier? What was the impact of a bed crisis and ward closures? What if patient flow into the Emergency Department was streamed – separating majors from minors?

The work helped us understand some key learning. First, if a 4 per cent annual increase in emergency admissions continued, both Acute Trusts would experience severe difficulties very soon. It could lead to 400 cancelled elective admissions per month after 5 years if no extra resources were available.

'The great virtue of modelling is that it is a safe place where we could try all sorts of ideas for improving things which would be totally impracticable to test for real.'

It became clear that the levels of GP referrals were a key factor in stabilising the system and that increased use of the Walk-in Centre was only moderately effective in reducing A&E workload. Of all the measures to reduce pressure on the system, preventing admission of older patients had the biggest impact.

The modelling of the Emergency Department demonstrated that streaming patients – dividing majors from minors - improved waiting times: especially for minor cases. However, there was a small increase in waits for medium severity patients, almost certainly avoidable in reality. There were benefits from using staff flexibly and responsively, driven by demand. The model could be used to develop rules for deciding when it was beneficial for the system to split patients into majors and minors streams.

The process we undertook had a number of things in its favour. It was the right model at the right time: it was dealing with a real world, urgent problem. The chair of the Steering Group was charismatic and enthusiastic about engaging everyone. It was important that people were collaborative and did not blame each other. The great virtue of modelling is that it is a safe place where we could try all sorts of ideas for improving things which would be totally impracticable to test for real, but, in the computer, took just a few seconds to run through a 5 year scenario. The model was simple and we could run it interactively with the stakeholders, which was appealing.

Our key messages to health and social care in Nottingham were that relatively small changes in one part of the system could have high impact on the whole system. It was clear that GP referrals were a key issue and that alternatives to admission were more effective in reducing bed occupancy than was discharge management. The model suggested that policy should focus on keeping less severe patients away from the ED and that there was a need for better outpatient services for diagnostics and treatment.

We presented our findings at a 'Stakeholder Day' at Nottingham Forest Football Club. The modelling helped in the development of a Local Services Framework and an Independent Sector Treatment Centre was opened whose goal was to reduce and better manage the flow of patients into hospital.

Sally Brailsford is Professor of Management Science, School of Management, University of Southampton.

> "It doesn't work to leap a twenty foot chasm in two ten-foot jumps."
>
> **American Proverb**

6. Clinical championship of modelling is vital in change management

When Torbay hospital considered closing some beds, clinicians and modellers teamed up to understand the impact. **Andrew Fordyce, Mike Williams and Michael Allen,** worked together to show the plan wouldn't work without better measures to reduce hospital length of stay.

Mr Andrew Fordyce is Clinical Systems Engineer in Torbay Hospital, where he is also a consultant oral and maxillofacial surgeon. He is looking at how to improve the safety, quality of experience and outcomes for patients across the health and social care system, aiming to create a reliable system 24/7. This should also reduce costs and reduce the number of hospital beds required.

Mr Fordyce has built a clinical-academic partnership combining modelling skills with expertise in patient safety, patient pathways and local clinical practice. Together, the partnership has looked at whether it is possible to ensure 24/7 reliability from the local healthcare system while closing some beds at Torbay Hospital. As a result of qualitative research and by modelling various scenarios, the partnership has concluded that it would be wrong to close beds immediately, until length of stay can be further reduced.

The team comprises Mr Fordyce and Dr Mike Williams, a former NHS chief executive, whose focus at the University of Exeter Business School is on whole systems in healthcare and how different parts interact. They are joined by the person who can model and test-drive different options and interventions that might be used: Dr Michael Allen, from the Peninsula Collaboration for Health Operation Research and Development, based at the University of Exeter Medical School.

'When we made this presentation, the clinicians said "That's our world" which gave the modelling and the academics credibility. The finance people understood the complexities and said: "We can't go ahead with the bed closures".'

In phase 1, analysing data and interviewing key staff, the team made important insights into large variations in demand for urgent care:

- Primary care: GPs, busy seeing appointments in the morning, tended to do home visits later in the day, hours after initial calls. Often they would then send the patient into hospital, but the call to ambulance, not being a 999 call, had lower priority, resulting in patients arriving in hospital late in the day, even though they rang in early. The patient's need was subordinated to local optimisation of parts of system – meeting 999 targets and GPs getting through their full list first at the surgery. As a result, while elective patients tended to arrive at hospital before 8am, sicker, more urgent patients arrived later, peaking in the afternoon, just as resourcing in the hospital was beginning to reduce.
- ED process inefficiencies: The Emergency Department reckoned that junior doctors could fully process one patient an hour in an 8 hour shift, while on the ground, because of process inefficiencies, the reality was actually much less - closer to one patient every two hours.
- Progressive increase and variability in emergency admissions: Though the increase in emergency admissions averaged just 1.6 per cent per year, managers often thought it was much higher, due to their concern at high levels of variability.
- High levels of hospital 'escalation'. At the time of the research, escalation was of three types: additional beds (but not properly staffed) on specialist medical wards, opening of a day case ward as an in-patient facility with a higher proportion of bank staff, and opening of an empty ward staffed by a temporary clinical team to care for the patients. The hospital was in escalation a large amount of the time, with medical wards in escalation almost all the time, more often than the surgical wards.

> "We allow our ignorance to prevail upon us and make us think we can survive alone, alone in patches, alone in groups, alone in races, even alone in genders."
>
> **Maya Angelou, poet**

Given these considerations, the team modelled and attempted to answer a business question: 'If we make some medical divisional changes aimed to reduce length of stay, can we reduce bed stock?'

'We need also to look at how we might smooth the flow of elective patients, which is more variable than emergency admissions for some hospitals.'

A dashboard of measures was built into the model, including possible scenarios for escalation, patients on trolleys and cancelation of elective procedures (resulting in a loss of income for the hospital). As a result of displaying the model to a senior group of managers and clinicians they concluded that within current processes, despite the addition of the divisional changes, it would be unwise to close beds. The model demonstrated to all that the changes would have too little effect on length of stay.

Some changes were suggested and agreed. 'We also realised,' explained Mr Fordyce, 'that we need a better balance between surgical and medical bed allocation so that medical wards are less often in escalation. We need also to look at how we might smooth the flow of elective patients, which is more variable than emergency admissions for some hospitals.'

Mr Fordyce commented: 'When we made this presentation, the clinicians said "That's our world" which gave the modelling and the academics credibility. The finance people understood the complexities and said: "We can't go ahead with the bed closures". The clinicians were all relieved and the exercise showed in a clear light to all present that we need to rethink our plans.

'The learning has made us think about new ways forward, such as shifting our wards from the current speciality-based model to one based on levels of dependency. We need to think about this because, under the present system, with medical wards on escalation, we have a lot of 'outliers' in inappropriate wards. Maybe, we need to think more in terms of how much care patients need rather than within which speciality their condition sits.

'We're testing weekend working, to smooth out some of the peaks and troughs. It makes Mondays much calmer. We are continuing to build the partnership with the NHS in Torbay and South Devon and with the University of Exeter and we are recruiting MSc and PhD students to help us answer and model the more difficult questions.'

7. How strategic modelling solutions for unscheduled care in Grimsby makes change safer and more effective

When commissioners wanted to reduce strain on beds in a hospital in the north east of England, modelling helped them understand the importance of diverting patients and identifying those with particularly long lengths of stay, explains **Peter Lacey** of The Whole Systems Partnership.

This is an example of how healthcare planners can use a modelling process called 'Systems Dynamics', combined with local engagement, to address key capacity and planning questions. It helps everyone to share insights and to feel confident to make changes that might otherwise feel more risky.

My goal at the Whole System Partnership is to help develop strategy around health and social care and support sustainable change. When modelling and simulation help these goals, then we use these techniques - which is quite often.

> *'We wanted to look at whether a service redesign could reduce A&E admissions and how some A&E patients might be diverted to the intermediate tier if, for example, a GP – someone competent – was put in place to divert patients to a more appropriate care setting.'*

We use an approach called 'Systems Dynamics'. This aims to understand the information and signals that managed systems feed back into themselves, producing different behaviours among those involved in the systems. The strengths are that systems thinking is holistic, avoiding the pitfalls of looking at just one part of the picture. It explores behaviour over time, taking in delays and feedback, getting to the heart of 'cause and effect' and system behaviour. It involves a lot of sharing of mental models and building models as a group. The whole exercise is rooted in 'thinking' rather just in 'processing data'. It is issue focussed and a learning tool that relies on iterative learning rather than achieving final answers.

Central to the exercise is a belief that co-production with local stakeholders and engagement with the issues provides the most likely route to sustainable long-term change. We therefore use Systems Dynamics to strengthen local partnership working rather than just to give the right answer – although we are pretty challenging when our analysis and modelling points away from the local consensus.

An example of the approach has been tackling unscheduled care and capacity challenges in a local district general hospital serving Grimsby and a surrounding rural population, with a total population of around 170,000 people. We have used the Systems Dynamics tool as a development process to look at the effects, for example, of introducing ambulatory emergency care. The process has helped raise lots of questions and improve understanding of the issues.

The hospital has around 440 beds, with A&E attendances at around 4,500 per month, of which approximately 1,000 are admitted. The Care Trust Plus (the equivalent of the PCT/CCG before April 2013) had already invested heavily in intermediate care services over recent years. A great strength for change – and for modelling – is that the area enjoys integrated health and social care commissioning.

The issue we have been tackling relates to the redesign of emergency ambulatory care, diversion of some A&E patients to an intermediate tier and a desire to achieve much earlier discharge of patients, particularly very long stayers, from hospital. We wanted to understand the likely impact on hospital bed requirements and also the impact on the intermediate tier and on community capacity.

> "The function of freedom is to free somebody else."
>
> **Toni Morrison, novelist**

In particular, we wanted to look at whether a service redesign could reduce A&E admissions and how some A&E patients might be diverted to the intermediate tier if, for example, a GP - someone competent– was put in place to divert patients to a more appropriate care setting. We wanted to know how this change might save on admissions. Also, we were asking what would happen if the hospital shifted to much earlier discharge, particularly for patients with a length of stay exceeding 28 days.

To get to the heart of the issues, it was vital to set these questions in the local context: the underlying demographics; recent trends in admissions and how they relate to changes in service design; historic levels of impact of the intermediate tier developments; the parallel development of an emergency ambulatory care model of service and any overlap or potential double-counting.

To drill down on what might be happening, we made some assumptions, which still needed testing and validation through an iterative process between professional/clinical opinion and data gathering. For example, we assumed that a third of medical admissions with potential for diversion into the intermediate tier would be diverted gradually over the following two years and that 40 per cent of these diverts would not need intermediate tier interventions. We assumed that the remainder receive Rapid Response support for three days; 25 per cent of these would then remain in a recuperation bed facility and a third of the other 75 per cent would be admitted to intermediate care (bed or home) whilst the other two thirds went home.

Around early discharge, we assumed that, in the cases of people currently staying over 28 days, they would be discharged 20 days earlier on average and all of them would go to a recuperation bed. For patients over 65 who were staying less than 28 days, their length of stay in hospital would fall on average by a day (from 7.25 days) and 10 per cent of this group would be transferred to intermediate care (bed or home).

Within the intermediate tier, we assumed that patients stayed on average in recuperation beds for 21 days. Intermediate Care referrals were divided roughly half and half between IC beds and being treated at home. Intermediate care at home or in a bed would be provided for six weeks (42 days). All of these assumptions were shared, refined and agreed as a reasonable scenario for modeling, although alternative assumptions could also be adopted thus enabling sensitivity testing of the model.

So what did we find out? Making these assumptions, we found, for example, in the medical wards that the 2009/10 hospital bed base of 146 beds (with 90 per cent occupancy) would have to rise to 161 unless the hospital invested in enhanced accident and emergency care or in the developments of intermediate tier pathways. However, if these services were put properly in place, the bed requirement for medical wards would fall to 96 by March 2014. We also found that a 'swing-bed' policy would need to be in place on these wards to accommodate up to 30 additional beds at peak times to avoid significant outliers being found beds outside the medical wards.

'Probably the biggest impact has been to give managers, clinicians and commissioners confidence … everyone really understood that siphoning off patients at crucial points can make a big difference in reducing strain on the system.'

With respect to the intermediate tier, we were able to enumerate in considerable detail the numbers of diversions and admissions to intermediate care and recuperation beds, provided the new pathways were optimised.

The information from this work has been helpful in lots of respects. It has helped thinking about: the effects of medical outliers and the demand for surgical beds; cancelled operations due to bed pressures; bed and ward structures for all medical specialties; what nursing and resources might be required for the new bed structure.

Probably the biggest impact has been to give managers, clinicians and commissioners confidence to develop GP provision to divert patients from A&E, to create intermediate care and to get the acute hospital to close some beds. Everyone really understood that siphoning off patients at crucial points can make a big difference in reducing strain on the system.

Peter Lacey is Director of
The Whole Systems Partnership
www.thewholesystempartnership.co.uk

8. Using simulation to transform the funding of care for people with long term conditions

Modelling the care of people with long-term conditions is suggesting exciting new funding mechanisms built around care needs rather than particular conditions, explains **Claire Cordeaux** of SIMUL8.

We are using simulation as a way to support the national Year of Care for Long Term Conditions programme. We aim to understand how people with long term conditions currently use services for the purpose of setting a Year of Care capitation fund for this group of patients. The simulation will allow the testing of "what if?" scenarios and will be available to organisations planning for implementation. It is hoped that this funding mechanisms could incentivise work across disciplinary and sector boundaries and allow providers and commissioners to work effectively together, aligning funding flows and patient need for support.

The potential dividends are considerable. Some 15m people live with long term conditions. That figure is increasing each year because of our ageing population. Their care is responsible for 70 per cent of NHS costs and they are a significant cause of ED and urgent admissions.

'What might be achieved if we planned care for people rather than for individual diseases? Can we differentiate groups of patients by need and costs to create an annual tariff and reduce emergencies and manage care out of hospital?'

The work has been built on research from the Scottish School of Primary Care. It shows that, if you have one long -term condition, you are likely to have more than one. Indeed, once a person is in their eighties, they could have up to eight long-term conditions. For example, some 82 per cent of people with COPD have at least one more long term condition. Likewise, only 14 per cent of patients with diabetes have just diabetes. And just five per cent of patients with dementia just have dementia. Yet, despite the fact that a multiplicity of morbidities is the typical situation for people with a long-term condition, they are often treated in different disease silos.

Our simulation aims to look at what might be achieved if we planned care for people rather than for individual diseases. We are exploring whether there are common patterns of service use for people with different long-term conditions. Can we differentiate groups of patients by need and costs to create an annual tariff? Can we work within that tariff to reduce emergencies and manage care out of hospital? Simulation helps, because it helps to ask better questions, to test assumptions and see if they are correct for different environments and sites.

Our goal is to create a simulation model. We will test it on seven pilot sites, so that we can come up with a national model that can be used locally. We are looking for common parameters that we can build into a model that can then be adapted locally.

The starting points are to stratify patients according to risk, assess their needs and assess what services they require. The idea is to look at the whole population and create a pathway based on need and services. The assumption is that good care comes from integrating health and social care teams.

So where are we now at the end of year two? We are able to calculate probability of access to services by long term condition groups. We can include changing states for patients. We know the numbers of long term conditions.

Simulation is performing different functions. It informs question development and data collection. It helps us work out what policy should be and how it can be implemented. It allows experimentation and hypothesis testing where no historic data is available. It enables research evidence to be applied to policy and practice development. It shares national assumptions meaningfully at a local level. It reduces risks in policy development by generating evidence for decisions. It makes change less of a leap of faith.

Claire Cordeaux is Executive Director, Healthcare for SIMUL8
www.simul8.com

> "Blame no-one, expect nothing, do something."
>
> **NY Giants locker room**

About the Cumberland Initiative

The Cumberland Initiative aims to build a healthier, wealthier UK by exploring new models of care which efficiently deliver better health outcomes. We're a not-for-profit organisation, seeking to update the NHS and stimulate a new UK industry delivering smarter process management in healthcare.

It's a network of communities – healthcare, business and academic experts. We're engineering a predictive NHS – one which anticipates, plans and acts effectively. We are applying systems thinking, engineering design, operational research, management science and statistics to demonstrate what can be done towards developing a more agile and efficient service in health and social care.

> If you limit your choices only to what seems possible or reasonable, you disconnect yourself from what you truly want, and all that is left is compromise.
>
> **Robert Fritz, composer**

TheCumberlandInitiative
TRANSFORMING HEALTHCARE QUALITY AND COSTS

Join us and help us to transform the NHS and build Britain a new industry.

For further information please contact:

- **Terry Young** on terry.young@brunel.ac.uk or
- **Elizabeth Deadman** on elizabeth.deadman@brunel.ac.uk

The Cumberland Initiative

c/o Elizabeth Deadman
Communications Manager (Healthcare)
School of Information Systems,
Computing and Mathematics
St. John's Building
Brunel University
Uxbridge UB8 3PH

www.cumberland-initiative.org