Influencing and Interpreting Health and Social Care Policy in the UK

Professor Eric Wolstenholme,

David Monk,

Douglas McKelvie,

and Gill Smith[1]

Abstract

Over the past three years the authors have used System Dynamics modelling to improve service delivery for patient pathways in a wide range of health and social care settings.

At the national level models have used to support the moderation of legislation by influencing national agencies and the upper house of parliament. At a local level models have been used to help health communities interpret legislation in a meaningful and shared way to achieve a more balanced and sustainable consensus for change. Here modelling was the means to review investment decisions from a "whole systems" and multiple agency perspective. Many such communities are now working towards consolidating the modelling process into regular planning activities.

This paper describes the background issues facing health and social care in the UK and then describes work carried out around the national issue of delayed discharge from hospitals and the local issues of commissioning services across whole health communities. Apart from describing the issues, model and example results, emphasis is placed on data and the role of models in designing data requirements for management of patient pathways. Particular reference is also made at the end to the process of application of system dynamics in health and social care services.

Introduction - Health and Social Care: the Structures, Drivers and Performance Dilemmas

Health and social care in England are at the centre of a modernisation agenda whereby the government sets out a programme of change and targets against which the public may judge improved services. The government has made reform of public services a key plank in its legislative programme and pressure to achieve these targets is therefore immense and can be "career-defining" for the heads of the various agencies concerned.

The modernisation agenda is rooted in the NHS Plan, a ten-year milestone plan for health and social care which was initially published in July 2000 and is revised and extended as each new planning period is entered. Figure 1 describes the scope of this ten-year plan, with additional details of government investments (see Figure 2) and targets (see Figure 3). For instance, a key initiative of the NHS Plan is to take the pressure off acute hospitals by providing new services, such as diagnostic and treatment centres and intermediate care. The latter issue is one of the key subjects of this chapter affecting both national and local thinking in the UK.

[1] Eric Wolstenholme and David Monk are Directors of Symmetric SD Ltd, (eric.wolstenholme@symmetricsd.co.uk). Douglas Mckelvie and Gill Smith are independent consultants.

Fig 1: NHS Plan 2000

- 10 year plan; Health Authority (HA) funding over 3 year cycle (not annually)
- New resources for intermediate care and Mental Health
- Implementation by Modernisation Agency - to be set up by autumn 2000
- A new relationship with the Department of Health (DH) based on subsidiarity (earned autonomy)
- New "concordat" with private sector (includes buying capacity to ensure wait times targets are met)
- Rationalisation of health regulation
- New processes for health complaints/suspensions
- More patient representation
- Appointments to boards no longer in hands of government
- NHS to buy capacity from private sector
- NHS Plus to sell occupational health services to employers
- NHS Lift to upgrade primary care buildings (cleaning, catering)
- National Treatment Agency (NTA) to pool resources for drug misuse
- Rationalisation of health training

Increased funding tied to targets. No change to status of Social Services - unless fail to achieve joint working (and can then be forced into a Care Trust with Health)

Fig 2: Investment

- £30M for clean-up of patient areas/Accident and Emergency (A&E)
- £5M development funds for trusts and HAs that are in top quartile ("earned autonomy") – total £500M. Under-performers will have to show improvement plans to earn their share
- £300M for equipment cancer, kidney and heart disease, including 50 Magnetic Resonance Imaging (MRI) scanners and 200 Computed Tomography (CT) scanners)
- 20 diagnostic and treatment centres by 2004 (day or short-stay surgery)
- 500 one-stop primary care centres
- £900M for intermediate care (hospitals to be kept for acute services) – includes 5000 rehab beds and 1700 non-residential rehab places
- 100 on-site nurseries in hospitals
- £300M by 2003/4 to introduce Mental Health service framework (in addition to £700M already announced)
- 50 early intervention teams to support young people and families
- 335 crisis teams
- 50 more assertive outreach teams over next 3 years (in addition to 170 planned for April 2001)
- 1000 graduate-level primary care workers to support General Practitioners and Primary Care Trusts (PCTs) and provide brief therapy (e.g. for depression)
- 300 staff to improve prison health services

Fig 3: NHS Targets

"A key message…..in formulating this Plan was that it needs a small focused set of targets to drive change. Too many targets simply overwhelm the service"

- By 2002, anyone with operation cancelled on the day for non-clinical reason will get new appointment in the same month or payment to go private
- By 2004 max wait times in Accident and Emergency 4 hours; to see GP 48 hours (or primary care professional within 24 hours)
- By 2004, 7000 extra beds (2100 in general/acute wards, 4900 in intermediate care)
- By 2005 – wait times for operations down to 6 months, outpatient appt to 3 months
- Reduced mortality rates for major diseases by 2010
- Reduced health inequalities (targets TBA 2001)
- Benchmarking cost of quality care (milestones 2003/4)
- By 2010, 100 new hospitals under Private Finance Initiatives (PFIs)
- Year on year improvement in patient satisfaction (including cleanliness and food)

Insert 1 provides updated guidance for the current planning period (2003-6). The emphasis is on growth (increased levels of facilities and staff) and transforming the patient experience (flexibility, choice, standards). It implies that there are known answers to current issues and all that is needed is focus, perseverance and attention to value for money.

Insert 1

The extra money coming into health and social services gives us the opportunity to make real improvements. We can expand through recruiting new staff, developing new services and creating new facilities. Even more importantly we can transform the quality of services by raising standards, tackling inequality, becoming more accessible and flexible and designing our services around the needs and choices of the people we serve.

This is about both quality and growth. The real test for success will be whether people can feel the difference and believe the services they receive are truly designed around them.

These are hugely ambitious goals. They will take time to deliver. Making progress over the next three years will be demanding and difficult and require real determination and discipline. It will need us to:

- Focus on priorities, we cannot make progress at the same pace in every area
- Extract the maximum value from every pound
- Be prepared to change old practices, be creative and take uncomfortable and difficult decisions in the drive to improve quality and respond to people using services

Priorities and Planning Framework 2003-6. Foreword by Nigel Crisp

In order to implement the modernisation programme, the government has set up various task forces and in order to ensure the new policies are providing the required results, the government has also strengthened the inspection and regulation framework and introduced national programmes.

The Potential for System Dynamics to Equip Ministers and Managers in Improving Health and Social Care

SD work in the UK health and social care field has been gaining momentum since the mid 1990s. Early work (Wolstenholme, 1993) laid the foundations for the creation of the concept of "whole systems thinking", a term which now has widespread use throughout the National Health Service (NHS).

Although the early manifestation of whole systems thinking was somewhat qualitative, it subsequently paved the way for more rigorous SD modeling (Wolstenholme, 1996 and 1999; Roysten, 1999, van Ackere, 1999; Lane, 2000; Dangerfield et al, 1999 and 2000; Wolstenholme et al, 2005; Lacey, 2005). To date the method is being extensively used by the economics and operational research section of the Department of Health as well as by a number of private health and social care consultants and academics. This UK work has also been complemented by other health related system dynamics work in other countries (Hirsch et al, 2005).

There is considerable evidence in this body of work and the presentation here, that health and social care managers are actively recognising the role of system dynamics for addressing the high degree of dynamic complexity and flux present in their organizations. In particular, the work is highlighting the need to improve in line with national guidelines and the interpretation of these at a local level.

There is evidence of a willingness to engage in the (often demanding) process of externalizing mental models and also a strong recognition that performance improvement will not come through collecting ever-increasing amounts of data. Indeed there is recognition that the modeling process of itself assists the definition of appropriate data.

However, in many important places there still persists a culture of change through redefining boundaries in isolation from processes and a focus on target lead managerialism, which contrast sharply with the aims of systems concepts (Figure 4)

Fig 4: Managerialism versus System Dynamics

Issue	***Managerialism***	***System Dynamics***
Defining the problem	May be taken as self-evident (as shown by “symptoms”)	Approached with caution: source of problem not seen as self-evident in complex systems
Defining the solution	Chosen as the course of action which appears the most cost-effective response to the problem	Chosen after experimenting with alternatives (checking likely responses in a complex system)
Type of solution	Framed in terms of management by objectives and targets	Focused on management of real operational processes
Reviewing outcomes	Tendency to simplistic cause/effect reasoning: e.g. X did not happen due to insufficient resources	Rigorous analysis of the behaviour of the whole system: X began to happen but caused a response in another part of the system….
Underlying principles	***Linear*** thinking, managers define actions and ensure ***compliance***, re-plan for next initiative. Tends to focus on organisation ***structure and boundaries***	***Systems thinking***, ***quantified*** view of dynamics and ***interdependency*** leading to (***innovative***) shared solutions devised by participants and ***iterative*** development of plans. Focus is on ***process and flows***

Influencing National Policy: Delayed Hospital Discharges

Issue

Delayed hospital discharge is a key issue which first came onto the legislative agenda in late 2001. The ‘reference mode’ of behaviour over time for this situation was that of increasing numbers of patients occupying hospital beds, although they had been declared “medically fit”. In March 2002, 4,258 people were “stuck” in hospital and some were staying a long time, pushing up the number of bed days and constituting significant lost capacity

The government’s approach to this issue was to find out who was supposed to “get the patients out” of acute hospitals and threaten them with ‘fines’ if they did not improve performance. This organisation proved to be social services who are located within the local government sector and who are responsible for a small but significant number of older people needing ex-hospital ('post-acute') care packages. Such patients are assessed and packages organised by hospital social workers.

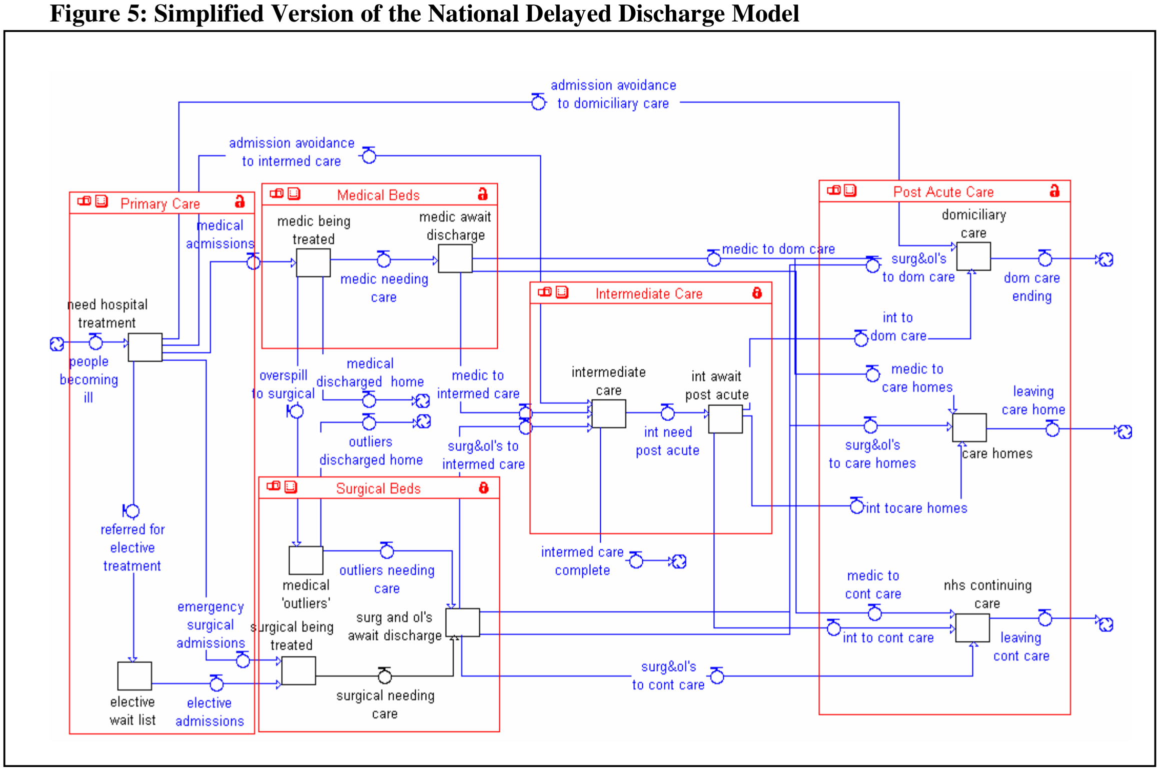
This approach was challenged by the Local Government Association (LGA-which represents the interests of all local government agencies at the national level) who suggested that a 'system' approach should be undertaken to look at the complex interaction of factors affecting delayed hospital discharges. This organisation, together with the NHS Confederation (the partner organisation representing the interests of the National Health Service organisations at a national level) then commissioned a system dynamics study to support their stance.

The remit was for consultants working with the representatives of the two organisations to create a system dynamics model of the 'whole patient pathway' extending upstream and downstream from the stock of people delayed in hospital, to identify and test other interventions affecting the issue.

Description of the Delayed Hospital Discharge Model

Figure 5 shows a highly simplified overview of the model created in the *ithink* system dynamics software. Complex behaviour over time arises in patient pathways by the interaction of the proportions of patients flowing (converging and diverging), their lengths of stays in different parts of the pathway and the capacities of each pathway, together with the actions taken by the agencies owning each piece of the pathway as patient movements conspire in and against their favour. The resultant behaviour of the whole system over time can be very counter intuitive, which is a major justification for the use of computer analysis to simplify the problem and improve understanding and communication. Each sector of Figure 5 will be described in turn.

Figure 5: Simplified Version of the National Delayed Discharge Model



Primary Care

This sector represents only patients who are becoming ill enough to be referred for hospital treatment. Patients here are classified by GPs into three main types: medical, emergency surgical and elective surgical. It is also possible for some admissions to be avoided or diverted. If space allows, a proportion of those needing admission can instead receive certain forms of intermediate care, or a domiciliary care package. This is normally at the expense of similar services available to a patient awaiting hospital discharge, but may be provided through "new money"..

Medical Beds

The model uses the shorthand term "beds" to mean essentially the total number of hospital places of a particular kind. In reality, hospital capacity is constrained by a number of factors, including availability of staff and operating theatre time. Similarly, the concept of hospital "capacity" in the model is simply a number, expressed in terms of so many beds. This is set at the acute hospital's intended level of occupancy (somewhere between 85% and 95% of the number of beds available).

In the model (but not shown in this diagram), there are various pathways through medical beds representing patients with differing characteristics (those with relatively straightforward conditions and few onward care needs, those with more serious conditions and who are more likely to have onward care needs, and elderly mentally ill patients, a significant sub-group of patients within medical beds, who have very specific onward care needs). These pathways cluster the medical admissions into groups, each having different characteristics with respect to length of stay, treatment time and assessment time. This clustering is important as a relatively small number of patients with complex conditions often require complicated care packages in post acute care.

There is a fixed capacity of medical beds. The model will only allow patients to be admitted if space is available. If medical beds are under pressure numerous coping strategies are brought into play to avoid over capacity problems. First, some patients are discharged home earlier than they ideally should be. In the short term, this creates some space for new admissions. However, the unintended consequence of this action is that some patients will need to be readmitted. Second, if there is spare surgical capacity, a number of medical admissions will be admitted to and treated in surgical beds, these are referred to as "medical outliers". Again, this has the effect of dealing with the immediate problem of patients needing to be admitted, but another unintended consequence is that elective operations may have to be cancelled resulting in increases of elective wait times. Mapping is a good way of surfacing these types of coping strategies and their unintended consequences.

Once patients have been treated, if they have no onward care needs, they go home. Those with onward care needs, and who therefore cannot be discharged unless a further service is available, go into the "await discharge" stock. They will remain in that stock, occupying a hospital bed, until the correct (intermediate or post-acute) service becomes available for them.

Surgical Beds

There are three kinds of surgical admission represented in the model.

Emergency surgical patients are always admitted, regardless of how much spare capacity exists. This means that at times the hospital will exceed its target occupancy rate, which resembles reality for the current users of the model.

Those needing elective surgery go on the waiting list. They are admitted as surgical beds become available. One of the buffers in the whole system, therefore, is the elective waiting list, which increases when the hospital is full, and reduces when the hospital has spare capacity.

Medical outliers are medical patients occupying surgical beds. These admissions take place only when there are no more medical beds available but there are surgical beds.

Patients are admitted to surgical beds in the model according to this order of priority: emergency patients (always admitted), medical outliers (if medical beds are full and there are spare surgical beds), and then elective patients (if there are spare surgical beds). As with medical beds, once patients have completed their treatment, they either go home, or, if in need of onward care, wait in an acute bed until a suitable resource is available.

Intermediate Care

In this simplified version of the model shown in Figure 1, there is one entity called "intermediate care". The actual model represents a number of services, some of which are intended as alternatives to admission, and others that assist timely discharge. Some intermediate care services do both. Intermediate care is time-limited. On completing a period of intermediate care, patients / service users either go home with no further service, or are discharged with a post-acute package. The latter might involve waiting in intermediate care until the post-acute care is available.

Post-Acute Care

There are three main types of post-acute care in the model, each with a separate capacity: domiciliary care, care homes (residential or nursing), and NHS continuing care. Patients are discharged to these services either directly from acute hospital or from intermediate care. Some are referred directly to the domiciliary care service as an alternative to acute admission.

Each post-acute service has a fixed capacity. Vacancies become available according to an average length of stay (care home or continuing care) or duration of package (domiciliary care). This figure includes those whose care package ends only on their death.

The Model and Data

As is often the case in developing system dynamics models the (extensive) categories of data that are currently collected by the organisations being modelled do not neatly match the data categories required by the model. Indeed the process of matching data and model structure can create significant insights and be a revelation to managers about the data they **really** need to manage their internal processes.

Modelling facilitates an in-depth focus on data. There are some data items that are simple "inputs" to the whole system, and which can be entered into the model as such. These would include:-

- The capacity of each service
- Average treatment or service lengths of stay
- Demand data at the entry point to the whole system (i.e. numbers of people requiring acute care)
- The percentage of people using a particular service who will go on to have a need for a further type of service

However, most data items are not in themselves "inputs" to the whole system. They are indicators of how the whole system is operating but they do not drive it. A useful learning point is that, even where agencies have accurate data for a particular part of the process, it is not always possible simply to put these numbers into the model. For example, the daily admission rate to intermediate care may be known, but it is not an input. It should be clear from the diagram that once the model is running, the rate of admissions to intermediate care is a product of the rate at which people are being referred to intermediate care (mainly) from the acute hospital, and the rate at which people are leaving intermediate care. And that rate is itself dependent on the availability of some post-acute services. The admission rate is therefore used, not as an input, but as a check that the model is correctly deriving this figure.

As the model is running, variables are checked to ensure their behaviour within the model corresponds to how they "really" behave.

In general, agencies hold more data about stocks (such as how many people there are in most categories at various points in time), when it would be more useful to know about flows (that is, the rates at which people move between different stages in a process).

The model reveals an absence of data that is absolutely critical for joint planning to improve the hospital discharge process. For example, there is very detailed information about "lengths of stay" in hospital for all categories of patient but the model shows that it is equally important to know:

- length of stay up to the point where a patient is deemed ready for discharge (which would be a model input called "treatment time"), and
- The length of time spent awaiting discharge (which would be a model output that would vary according to whether there is space available in a given post-acute service)

Other useful data (currently unavailable) are the proportions of patients being discharged to the main post-acute options of domiciliary care, care homes and NHS continuing care. These categories do not neatly map against the standard discharge data available.

In the case of the delayed discharge model the best available national data was assembled.

Configuration of the model

The model was set up to simulate a typical sample health economy over a 3 year period when driven by a variable demand (including three winter "peaks"). The capacity constrained sectors of the model were given barely sufficient capacity to cope. This situation was designed to created shocks against which to test alternative policies for performance improvement. Major performance measures in use in the various agencies were incorporated. These included:

1. cumulative episodes of elective surgery
2. elective wait list size and wait time
3. numbers of patients in hospital having completed treatment and assessment, but not yet discharged (delayed discharges)
4. number of 'outliers'

The model was initially set up with a number of fixed runs, to introduce people to the range of experiments that yielded useful insights into the behaviour of the whole system. From there, they were encouraged to devise their own runs and develop their own theories of useful interventions and commissioning strategies.

The three main polices tested in the fixed runs were:

1. Adding additional acute hospital bed capacity. This is the classic response used over many years by governments throughout the world to solve any patient pathways problem and was a favourite 'solution' here.
2. Adding additional post acute capacity, both nursing and residential home beds but also more domiciliary capacity.
3. Diverting more people away from hospital admission by use of pre-hospital intermediate capacity and also expansion of treatment in primary care GP surgeries.

Example results from the delayed hospital discharge model

Figures 6, 7 and 8 show some typical outputs for the delayed hospital discharge model. Figure 6 captures the way capacity utilisation was displayed (actual beds occupied v total available for both medical and surgical sectors of the hospital) and shows the occurrence of 'outliers' (transfers of patients from medical to surgical beds) whenever medical capacity was reached.

Figures 7 and 8 show comparative graphs of 3 policy runs for 2 major performance measures for 2 sectors of the patient pathway - delayed discharges for post acute social services and cumulative elective procedures for acute hospitals. In each case the base run is line 1. Line 2 shows the effect of increasing hospital beds by 10% and line 3 shows the effect of increasing post acute capacity by 10%.

The interesting feature of this example output is that the cheaper option of increasing post acute capacity gives lower delayed discharges and higher elective operations whereas the more expensive option of increasing acute hospital beds benefits the hospital but makes delayed discharges worse. The key to this counter intuitive effect is that increasing post acute capacity results in higher hospital discharges which in turn reduces the need for the 'outlier' coping policy in the hospital, hence freeing up surgical capacity for elective operations.

Figure 6. Medical and surgical bed utilisations in hospital and 'outliers'

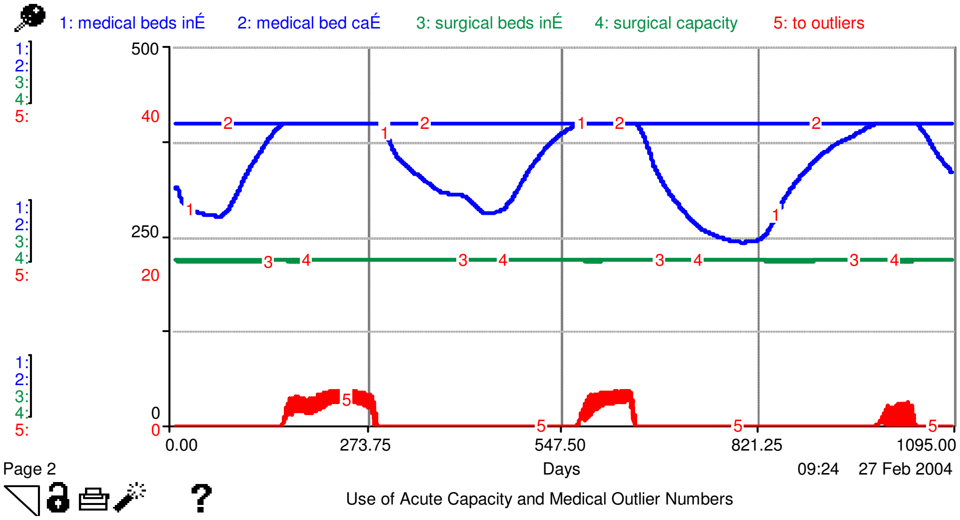


Fig 7.**Delayed hospital discharges for 3 policy runs of the model.**

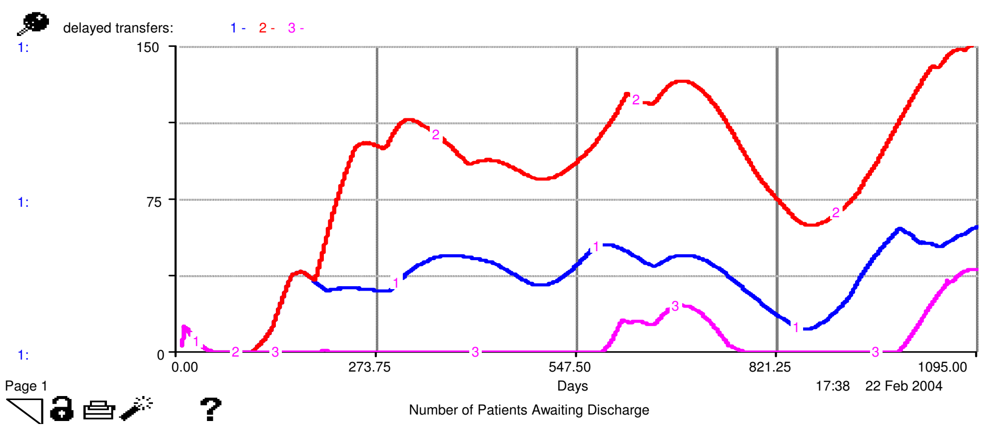
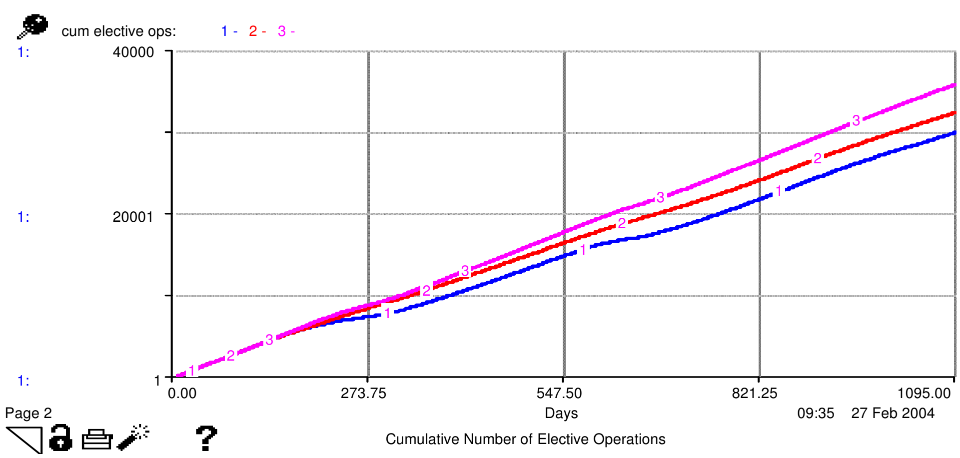


Fig 8 **Cumulative elective operations for 3 policy runs of the model.**



Figures 7 and 8. Delayed hospital discharges and cumulative elective operations for 3 policy runs of the model.

Conclusions from the delayed hospital discharge model

Common sense solutions can be misleading

The obvious unilateral solution of adding more acute capacity can exacerbate the delayed discharge situation. More hospital admission will be made in the short term, but no additional discharges will be possible. Hence, the new capacity will simply fill up and then more early discharges and outliers will be needed.

Fines may have unintended consequences

This solution will also be worse if the money levied from social services is given to the acute sector to finance additional capacity. It will be worse still if it causes the post-acute sector to cut services. The effects of service cuts may also spill over into other areas of local government including housing and education.

There are some interventions that can help

1. Increasing post acute capacity gives a win-win solution to both health and social care because it increases all acute and post acute sector performance measures. Further, counter intuitively, increasing medical capacity in hospital is more effective than increasing surgical capacity for reducing elective wait times. This again is because 'outliers' are avoided.
2. Reducing assessment times and lengths of stay in all sectors is beneficial to all performance measures, as is reducing variation in flows, particularly loops like re-admission rates
3. Increasing diversion from hospitals into pre-admission intermediate care was almost as beneficial as increasing post acute capacity.
4. If fines are levied they need to be re-invested from a whole systems perspective. This means re-balancing resources across all the sectors (NOT just adding to hospital capacity).
5. In general the model showed that keeping people out of hospital is more effective than trying to get them out faster. This is compounded by the fact that in-patients are more prone to infections so the longer patients are in hospital, the longer they will be in hospital
6. Improving the quality of data is paramount to realising the benefits of all policies.

An interesting generalisation of the findings was that increasing stock variables where demand is rising (such as adding capacity) is an expensive and unsustainable solution. Whereas increasing rate variables, by reducing delays and lengths of stay, is cheaper and sustainable.

Impact of the delayed hospital discharge model

This model was shown at the Labour Party Conference of 2002 and generated considerable interest. It was apparently instrumental in causing some re-thinking of the intended legislation, so that social services Departments were provided with investment funding to address capacity issues, and the "fines" (re-titled "re-imbursement") were delayed for a year. Reference to the model was made in the House of Lords (see insert 2).

Insert 2

"Moving the main amendment, Liberal Democrat health spokesperson Lord Clement-Jones asked the House to agree that the Bill failed to tackle the causes of delayed discharges and would create perverse incentives which would undermine joint working between local authorities and the NHS and distort priorities for care of elderly people by placing the requirement to meet discharge targets ahead of measures to avoid hospital admission….. ***He referred to "ithink", the whole systems approach being put forward by the Local Government Association, health service managers and social services directors involving joint local protocols and local action plans prepared in co-operation"***

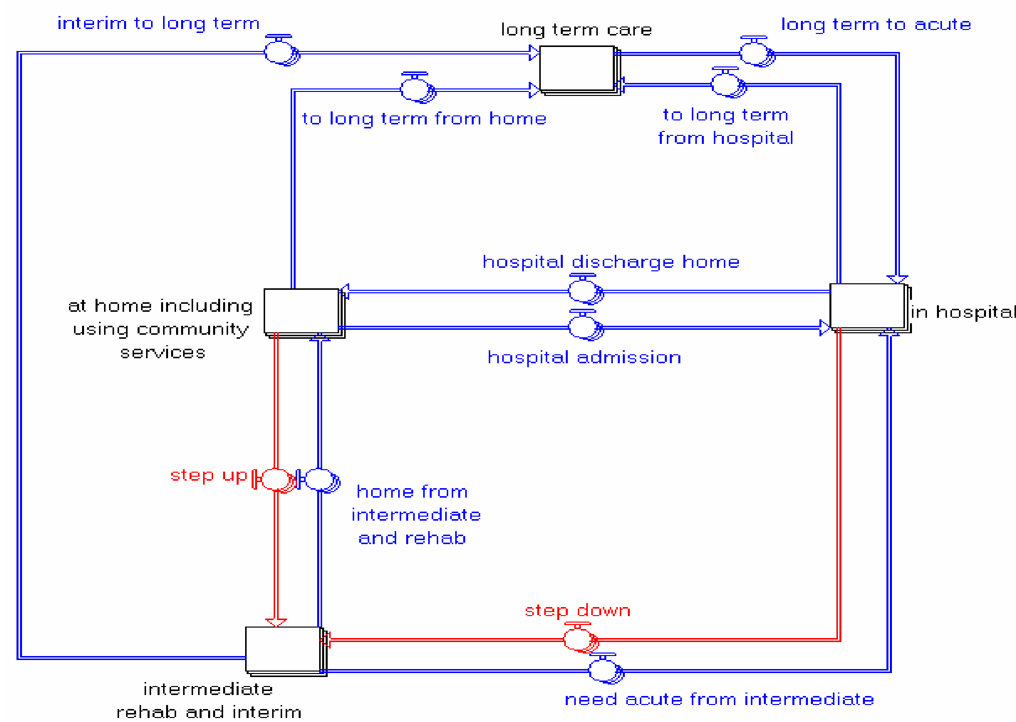
Interpreting national policy – use of patient pathway models for local commissioning and capacity planning.

Issue

As a result of the success of the national work, a number of local health communities (combined primary care trusts, acute hospitals and social services departments) were interested in developing similar models on a local basis, specifically for their own delayed discharge issues but more generally for commissioning (buying of services) decisions and capacity planning.

By definition this work required a model of significantly different configuration from the national delayed discharge model, which was limited to flows in and out of hospital. Local health economies are more complicated than this, with flows into social services from both hospital and from the community. However, in each locality it was found useful to use the national model as a starting template to communicate ideas. A generalised version of the emergent local health and social care economy model is shown in Figure 9 in the *ithink* system dynamics software.

Figure 9. Overview of the local health community model



The main differences between the new model compared with the delayed discharge model were:-

- the focus on older people rather than the general population
- the need for the model to focus on the population of a health economy, rather than based around the users of one district general hospital
- the need to represent a wider range of hospital beds than just medical and surgical
- the need to take account of all older people using social care services (mainly care packages at home, and residential / nursing home care) in that locality and not just those referred from the hospital
- the complexity of intermediate care and its location in the middle of a web of services

- the need to represent delays within services as well as those resulting from people waiting for a place in the "next" service

The altered shape of the diagram depicts a more circular representation of people flowing around a system, rather than the more left-to-right direction portrayed in the national work. Not all of the differences are obvious from comparing the shapes of the respective diagrams. The main difference is that the symbols in the new diagram look more three-dimensional, denoting that a number of similar flows are "arrayed" on top of each other. These arrays have two dimensions, people (divided into four age bands), and types of service (different for hospital, intermediate and Long Term parts).

.The age bands adopted were:

- adults aged 18 – 64
- older people aged
    - 65 – 74
    - 75 – 84
    - 85 and over

Apart from differences in structure of the local models, there were also significant differences in the style of projects for local application. The main differences were:

- A wider range of issues to be addressed
- More detailed and dynamic complexity to be modelled
- Greater size and variation in members of the management teams
- Wider range of people to report to
- More data quality issues
- A range of possible approaches by which users access the model

Description of Model Sectors (Service Types)

The classification of services in the local modelling was agreed to be of "service type" rather than service name. These were:

At Home

Most people in the model are "at home" and a subset of these is using community services (though the main service type represented in this version is people receiving social services-funded packages of care).

Hospital

Based on an incidence rate (proportion of people becoming unwell per day), every day a number of people are referred to hospital. Although not included in the diagram; the actual model represents the stages of presentation at Accident and Emergency and further assessment (including overnight) in an acute admissions unit.
There are seven main routes through hospital.

- medical patients (mainly emergency admissions)
- patients in surgical beds, subdivided into:
    - elective patients
    - emergency patients
    - medical "outliers" occupying surgical beds
- older people's medicine
- older people's mental health
- patients in hospital beds other than the local general hospital

As an alternative to referral to hospital, the model also shows the use of intermediate care as a means of admission avoidance (labelled stepping down in the diagram).

Intermediate Care, Rehab, and Interim Care

The model represents intermediate, interim and hospital rehabilitation services as an apparently similar group of services, but each has its own particular features with different flows through (which cannot be deduced from the summary diagram). The common feature of these services is their more transitional nature.

Intermediate care is a range of services that are between home (frequently combined with a care package) and hospital. People are either admitted to intermediate care as an alternative to going into hospital, or as a stage in their discharge from hospital services. Intermediate care tends to be time-limited, for up to six weeks, but this policy can be relaxed.

Some intermediate care services provide for only one option (either admission avoidance or assisted discharge). Others can be used by people referred from home or hospital. Some intermediate care services are provided in someone's own home (including supported housing), and others are bed-based (either in care homes or in hospitals). The desired outcome from a period of intermediate care is that someone will return to full independence at home, or will manage at home with assistance from the mainstream (i.e. not intermediate care) services.

Intermediate care is closely related to hospital rehabilitation, and to older people's medicine. In the model, older people's medicine is represented in the "hospital" part, whereas hospital rehabilitation is represented in the intermediate, rehab and interim part. For very good reasons, intermediate care services often have the word "rehabilitation" in their title.

Hospital rehabilitation, in the model, is similarly located between hospital and home, but the flow is only in one direction, from hospital to home. Where it differs from intermediate care is that:-

- it is less focused on frail older people
- it is not time-limited
- it is not used for admission avoidance – it is a service for people whose recovery from an acute episode requires a therapeutic element

Long Term Care (post acute)

There are five service types in the Long Term care part of the model:-

- residential care
- nursing homes
- residential care (mental health)
- nursing homes (mental health)
- NHS continuing care

Long term care means a form of care that becomes a person's permanent home. Although people do return home from a long term placement, these numbers are small enough that they do not need to be represented in the model. However, it is possible to move between the different types of long term care in the model. It is also possible for people whose main location is in long term care to be admitted to hospital.

People flow into long term care either directly from home, from hospital, or from interim care. In some boroughs there are policy aspirations that might exclude some of these movements, e.g. people should not be admitted to a long term placement directly from hospital.

Obtaining Data for - and Learning about Data from - the Model

In a modelling project of this scale it is even more inevitable than in the national case that agencies will find that they do not have data available in a form that matches the model structure. This is because it is unlikely that any standard databases have been derived from a need to represent the stock / flow structures of a whole system of services.

Taking a whole systems perspective, the challenge is to understand how particular population groups (older people with differing degrees of dependency needs) move around between different service levels and types. This requires information about how many there are (stocks), where they are currently (stocks, but best reported over time), and the factors that govern how they move into, out of, and between service types (flows).

The "currency" of data is also important, and whether it is being used for management information or performance reporting. For example, in reporting performance, a typical measure might be the number that have used a particular service in the course of a year. This is a stock-like measure, but if reported in the absence of measures of turnover is much less useful than a daily (or even weekly) measure of "the number using service at this time", ideally plotted on a graph over time.

Data relating to discharge from hospital exemplify many of the differences between the kind of information that would be useful to managers (as set out in the model) and the kind of information that is actually collected. This is particularly pertinent because hospitals represent the services that have the highest throughput, and are probably most susceptible to becoming "blocked" as a consequence of the current state of other services.

There is a standard national set of data codes (although with some local variation in exactly which codes are used). Clearly, these codes represent categories of information that were once believed to be important, perhaps for accounting purposes. Sadly, they do not shed much light on the effectiveness of the whole system of services for older people. For example, the majority of people are discharged to "home / usual place of residence". Buried within that category are:

- People who returned home with no need for any further service
- People who were referred for a care assessment, but received no service (a significant number as it turns out)
- People who were discharged home with a new care package

Additional problems arise where people whose normal place of residence is a care home, from which they were admitted to hospital, are recorded in this category (because they returned to their usual place of residence). We thus "miss" a proportion being discharged to a care home. The various categories of care home in the coding framework are also frequently not up-to-date.

In essence the codes used by social service about where people come from are quite different from the codes used by hospitals as to where people go to.

In order to overcome some of these difficulties, many people worked extremely hard to extract the most meaningful data from their respective systems. Further, data research was carried out and spreadsheets were developed to transform the officially reported data into the model categories by allowing a user to vary a number of assumptions about what is going on within each coded category.

Model configuration and access

The model was set up to simulate time-steps of one day for five years and most model outputs were presented in the form of time-series graphs. Because the model was detailed and was designed to address a wide range of issues it had numerous policy levers and data input screens. Consequently it was relatively slow to run simulations and required considerable thought about the best medium by which to allow users access to the model

It should be emphasised that the main purpose of this type of model is as a learning environment within which a group jointly investigates the behaviour of a complex system. The model is not a backroom tool that is fed data and gives some answers (for example, how many of these will there be in two years time?) The power of the model is in its representation of the system's behaviour rather than in the exact figures that it produces. It is useful to know that:

- if this service gets blocked some other service will suffer
- there always seem to be a number of delays within this service but they do not impinge on other parts of the system – perhaps we can live with this
- if we divert people from here to there, this will have unintended consequences (in the form of other people who also needed to go there being blocked or having their progress slowed down)
- this kind of delay is more significant than some other kind
- if we change this capacity, we also need to change that one and monitor this other thing closely

- there is a tendency for this service to be (counter-intuitively) under-utilised when the whole system is under pressure

The optimum way to use this tool was found to be to allow users to access the model in facilitated workshops or group events, making various assumptions about how key service inputs might change over a period of five years. The model was set-up to pause every quarter, at which point it was possible to make some additional changes to the model settings (such as add or take-away capacity – open up some new services or readjust some service priorities) and then continue the run.

The most common model settings, applicable to most service sectors, were:

- The length of stay (with a wider range of settings covering the different stages in hospital) and percentage discharged to (different destinations depending on service type) settings
- Service capacity
- Allocation policies (when a service is under pressure, how admissions are prioritised based on current location of persons being referred)
- Availability and workload size of care managers
- Population (including demographic changes over the five year period) and incidence rates (e.g. rate of referral to hospital, or for community services, of people accessing from the community)
- Assumptions about the demand for elective procedures
- Extent to which intermediate care is being prioritised to assist discharge or admission avoidance

The model then displays the results of these inputs and changes, in the form of a series of graphs. Mostly these graphs show the changes in the values of stocks (or sums of stocks) over time. It is also possible to plot graphs of flows, but in a model of this complexity these tend to be used more for technical (validation or testing of the model) purposes.

Examining some Model Outputs from Scenario Meetings

Model Outputs

Scenario planning meetings were dynamic, in that they consisted of a group of stakeholders inter-acting with the model, facilitated by the consultant. These meetings produced interruptions, “what-if?” questions, and discussions about “why did that happen?” in a relatively unstructured way.

It is difficult to capture the nuances of such an event in words. However, Figure 8 shows some examples of how the model and the workshops were used to generate insights. All graphs show changes in stock levels over a five year period measured in days.

The graphs in the top left quadrant show:

1. the numbers of people actually receiving care packages at home (dom total) compared with the capacity for this service (dom capacity),
2. the level of care management in hospital (hospsw cap) mainly dealing with patient assessment for discharge (this is a function of number of care managers multiplied by average caseload), again compared with the capacity of the service.

The graphs in the top right quadrant show:

1. the daily numbers of people in hospital accident and emergency (A&E) who are deemed to need medical or older people’s medical beds
2. the total number of residents being screened in A&E per day.

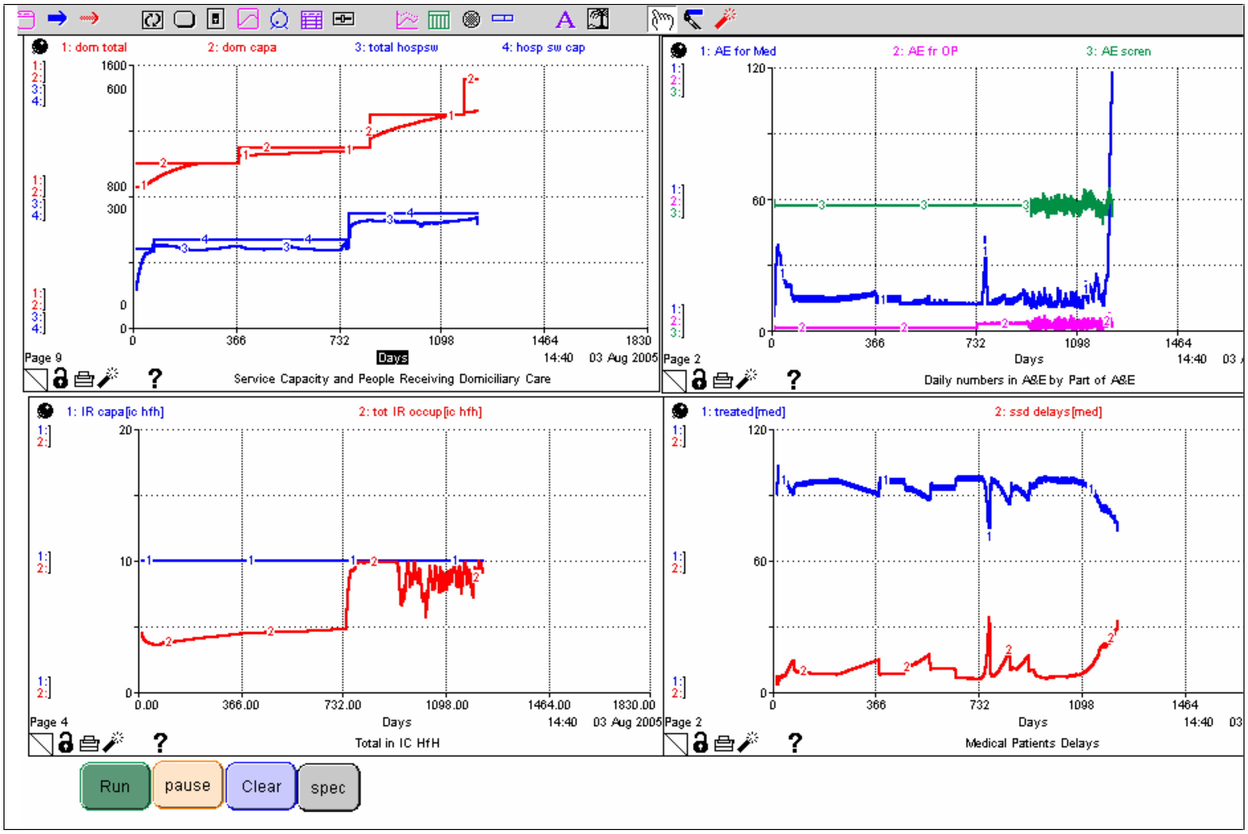
The graphs in the lower left quadrant show:

the numbers of people in one particular domiciliary-based scheme for intermediate care, again compared with the capacity of that service.

The graphs in the lower right quadrant show:

The number of people being treated in hospital because they need to be there) and those delayed (for social services assessment or awaiting service) in medical beds.

Figure 8 Example results from a Scenario Workshop using the Local Health Community Model



Experiments over the first 2 years of the run (days 0- 730)

The group first tackled the interaction between social work discharge assessment in hospital and delays in hospital accident and emergency rooms and in hospital medical treatment.

Early in the run (top left) it becomes apparent that hospital social work capacity is going to be stretched. Note the sharp rise in A and E medical delays around that time (top right) as delays occur in older people in hospital waiting for assessment to be discharged (bottom right).

As an immediate measure, an extra hospital social worker caseload is added. Immediately, the number of delays in hospital drops (but does not go away completely). Also, the potentially worrying performance in A&E stabilises. There is still some spare capacity in care packages, but not much.

Around days 60 – 200, there are still some social service delays in medical beds, even although there is spare capacity in care packages and among care managers. Note that the number of delays remains fairly constant. This is an example of what might be termed a “normal” level of delays. If referrals are typically made, say, two days before patients are deemed ready for discharge, and if the process of allocation to a care manager plus time taken to complete the assessment come to more than two days, there will always be a number of people being assessed for care beyond their statutory discharge date. This is a problem that cannot be solved by changing capacities, but only by ensuring that referrals for care assessment are made earlier in an admission, or by (somehow) speeding up assessment time. This shows the benefit of presenting data about delays as a time-series graph. Generally, delays that remain at a certain level, neither rising nor falling, are less worrying than delays that rise steadily over time.

Just after day 200, delays in hospital begin to do just that. Note that the delay line moves from being horizontal to rising just as the existing capacity of domicilliary care packages is reached. It goes on rising until the end of the year, at which point additional care packages are put into the system, and hospital delays revert back to their “normal” level. Similarly, around day 550, delays have started to rise once more. This is due to pressure on residential and nursing home places (graph not shown on this display), a problem that is resolved by increasing their respective capacities.

Experiments over the third year of the run (days 730-900)

At the end of the second year, the group tackled a different issue, which was to consider why the intermediate care service (bottom left) was relatively underused.

This was a service that mainly took referrals from people at home (and at risk of admission to hospital), people currently in A&E, and people in assessment wards. By resetting the hospital admission avoidance variables, making admission avoidance available to more people who would otherwise be admitted, this small intermediate care service quickly fills up.

But note the effect that this has on A&E (top right) and hospital treatment and delay (bottom right). With more hospital admissions being avoided, one would expect A&E throughput to increase and hospital delays to reduce. In fact, the opposite happens. Delays increase rapidly, meaning, among other things, that admissions from A&E take longer, and people begin to build up in A&E.

This turns out to have been caused by an increase in the assessment load of hospital social workers who also have to do pre hospital intermediate care assessment in preference to hospital discharge assessment. Once care managers reach full capacity, assessment gate-keeping for a range of services slows down. This causes more delays within the hospital which then feedback into A&E, the very service that this measure was intended to help.

An increase in care managers around day 730 resolves this problem. The main lesson here is that increasing service capacity without considering gate-keeping arrangements can have unintended consequences.

Experiments after the third year of the run (days 900 -)

From day 900 onwards, this group tried various other experiments with the model. The periodic rises in hospital delays were caused by further shortages in care home places (and resolved by increasing these resources). The less even graph-lines from day 930 onwards result from some experiments being done to introduce a daily variation (though the same average) to the rate of referral to A&E. Towards the end of the run, the sharp increase in hospital delays is, once again, due to a lack of assessment resources. It is in runs like this that the true usefulness of the demographic inputs is apparent. There is a gradual increase in the oldest population group over time, with a proportionate effect on demand, but no compensating increase in post-acute services.

Common Learning Points from Scenario Meetings

Types of delay:

Normal or Unavoidable Delays

Some delays just have to be expected. An example is delay in hospital discharge. The average time taken to complete a discharge assessment and provide post hospital care is simply often longer than the average time between notification of discharge and actual discharge date. Hence some delay occurs not as a result of blockages but because procedures do not quite match-up. This kind of delay will normally take the form of a straight line when plotted on a time-series graph. Such delays do not build up and, if there is sufficient capacity in the holding service, normally hospital, it is hard to consider them a problem.

Delays where Services are at Capacity

If a service is full, then the maximum number of new admissions per day is equal to the number of existing users leaving it per day. Any imbalance between arrivals and departures will result in increased numbers of people in the service and delays. The typical shape when plotted on a time-series graph is a

line that is rising, with the steepness of the gradient reflecting the size of the difference between turnover and new referrals.
Obviously these delays are more serious than the "normal" type. If the number delayed within a service is rising, eventually this backlog will affect the operation of that service and contribute to backlogs further back in the whole system.

Delays where Assessment Teams are at Capacity

These delays are similar to those arising from service delays, but can be more serious with a broader and more pronounced effect. This is because assessment teams cover a range of services affecting a range of patient pathways. Hence a problem in one location can have a serious 'knock on' effect on services in many locations.

Further, the cause of a blockage in an assessment process might well not be due to change in the capacity of that particular service. The cause of the blockage might be that some other specific service has become blocked (or slowed down), thereby over-stretching the assessment team.

Even worse, this interplay between delays often leads to the perverse underutilisation of certain services when the whole system is otherwise stretched. This phenomenon is most observable in intermediate services, because they are taking service users in the middle of a patient pathway.

Intermediate Care

Most of the local groups were interested in modelling the dynamics of intermediate care services, which are seen as a major solution to reducing bottlenecks in patient flows. Modelling certainly shows why it can be hard to plan such services. They are sensitive to what is happening in just about every other part of the system. Moreover, they tend to have very small capacities relative to the size of the other services within the whole system, and are typically sub-divided into several schemes such as bed-based and community-based, hospital admission avoidance and timely hospital discharge.

The model was created with a very elaborate set of menus covering not only "who this intermediate service is for" but also how to prioritise between people coming from many different destinations, overlaid by four age bands.

A consistent leaning point was that small changes in priority settings have large effects. Because the services have a low capacity, expanding the admission criteria can lead to them being immediately swamped with new referrals. The services are also intended to be time-limited, but that cannot prevent people from becoming "blocked" within intermediate care, given that a significant proportion of those in intermediate care can only leave if a domiciliary care package is available.

Conclusions from creating commissioning models for local health communities

Despite the challenges of having to model in more detail, with variable quality data and a wider range of issues across a management team of more variable constitution, the usefulness of describing services in basic inflow-stock-outflow terms, together with seeing the interconnectivity between such descriptions and behaviour over time, turned out to be particularly revealing. Benefits included:

1. Allowing risk-free testing of the effects of policies across patient pathways and an operational focus on ways to reduce costs.

2. Helping key individuals in different agencies developed an increasing awareness of their role in respect of the whole patient pathway.

3. Developing understanding of the way patient pathways worked by:

    - Improving the definition of service capacity and its relationship to referral rates, length of stay.

- Explicitly defining the component states of being in hospital services as treatment, assessment and waiting for discharge.
- Introducing the concept of equilibrium. Whilst people expected that in most services at most times the numbers being referred and accepted for service are balanced by the numbers leaving, they were often surprised at the degree of instability created by small changes, particularly to policies around intermediate care.

4. Identifying inconsistencies between maps of the way the patient pathways were claimed to work and the data collected for the pathways.

5. Understanding the most useful configuration of people and services to enhance performance. Whilst a breakdown of patients by age was used in the modelling, there was a continual discussion about the alternatives to this, such as a breakdown by 'degree of dependency'.

6. Defining the data required for patient pathway management across agencies. One major example was acute hospital data discharge codes being unmatched to the admission codes used for admissions to post acute services.

7. Facilitating discussion on defining and categorising issues and on performance measures.

Implications of the work for the process of carrying out system dynamics projects –the challenge of changing thinking.

System dynamics is most beneficial when, as described here, it is applied to deep-seated issues concerning long processes, which across multiple agency boundaries (Wolstenholme, 2003). Such projects are not only technically challenging but highly emotive and political. Users of the method need to recognise the context as well as content of project,

Technically, managers are unaccustomed to the fundamental nature of the questions posed by system dynamics. Questions such as "Is this care pathway complete/correct?", "Can people go anywhere else?", "What assumptions are built into routing?" and "How long do people spend in each stage?" are more fundamental than the usual ones they face every day. Further, managers do not tend to think in terms of 'flows'. They may assert that they need more resources, but often do not justify the request with an argument based on flows. Politically, multi-agency teams are also driven by professional allegiances and there can be conflicts of time and interest amongst the participants.

In order to get the best out of a system dynamics study it is important to have:

1. well-defined issues
2. consistent and representative teams of management professionals from all agencies involved, well briefed in the principles of the approach
3. well-defined expectations and well-scoped models in time and space
4. mechanisms for managing the politics of the management team and keeping senior people in touch with the project.
5. ways of maintaining momentum
6. ways of reinforcing the benefits

The Symmetric approach to system dynamics used throughout the work addresses the above issues by:

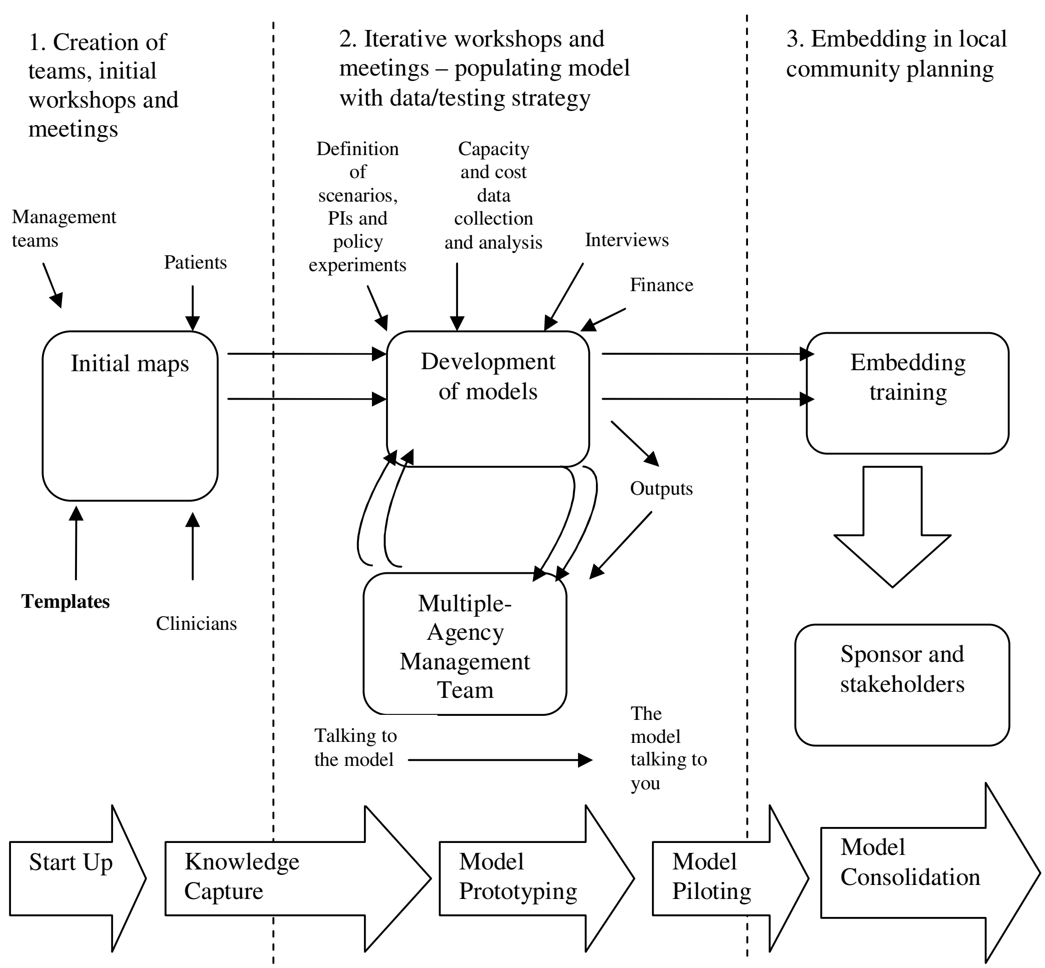
- An emphasis on using modeling to create environments for learning: participative approach and involvement of the management team at every stage. No "backroom" tours de force, presenting a completed model to an astonished audience!
- A willingness to combine the use of relevant national (generic) templates with a flexible approach to the specific needs associated with local applications
- Careful discussion of model development timescales
- A commitment (however time-consuming) to help local projects with the issue of data mining. It requires significant effort to find the requisite data to animate the model. Organisations typically

measure "snapshot" data (e.g. current usage of beds) rather than flows and participants may need to estimate proportions flowing through each route as well as lengths of stay

Figure 11 captures the major elements of the participative stages advocated here for successful system dynamics projects in multi-agency situations. It needs to be made clear from the outset that the work will be challenging. Successful projects require:

- partners in a multi-agency context to declare their assumptions and share their mental models of how the "system" works. There is no room for narrow views or simplistic apportionment of "blame". Participants must be able to "stand back" from single agency issues and see the broader patterns of the whole system (causal loop mapping)
- partners to agree the key issues they want to elucidate and address. These must be of significance to the behaviour of the whole system. It is important that they recognise the interdependence of their agencies, so that no policy or practice is considered "private" or "non-negotiable"
- painstaking work to unravel the structure (process) in a complex system and agree how to represent it in stock-flow terms, to achieve the simplest representation which reflects current realities
- a recognition that models are never complete. There is always more to do.

Fig 11: Stages of a System Dynamics Project – a participative approach



Overall Conclusions

This chapter has described recent and continuing work carried out by the authors in health and social care in the UK and some guidelines for improving the success of such projects. This work has been at the forefront of both national and local policy analysis and has achieving outstanding success in influencing behaviour and decisions by its timeliness and relevance. Many managers and clinicians have been involved in sponsoring and participating in the work. These people have proved very receptive to the ideas and shown a willingness to embrace the technical and political challenges of using the ideas in a comprehensive way. They have also been very appreciative of the benefits.

Acknowledgements

The authors wish to acknowledge the help and cooperation from many health and social care clinicians and managers who contributed to the studies reported in this chapter. The results reported and views expressed for the purpose of this chapter are entirely those of the authors and are not necessarily the views of the organisations involved.

References

Ann van Ackere (1999), *Towards a macro model of Health Service Waiting Lists*, System Dynamics Review, Vol. 15, No 3.

Dangerfield, B. and Roberts, C. (1999), *Optimisation as a statistical estimation tool: an example in estimating the AIDS treatment free incubation period distribution*, System Dynamics Review, Vol. 15, No. 3.

Dangerfield, B., Fang, Y and Roberts, C. (2001), *Model based scenarios for the epidemiology of HIV/AIDS: the consequences of highly active antiretroviral therapy, S*ystem Dynamics Review, Vol. 17, No. 2.

Hirsch G., Homer J., McDonnell G. and Milstein B. (2005) *Achieving Health Care Reform in the United States: Towards a Whole System Understanding,* Paper presented at the International System Dynamics Conference, Boston, USA.

Lacey, P. (2005) *Futures through the Eyes of a Health System Simulator*, Paper presented to the System Dynamics Conference, Boston, USA.

Lane, D. C., Monefeldt, C. and Rosenhead J. V. (2000) *Looking in the Wrong Place for Healthcare Improvements: A system dynamics study of an accident and emergency department,* Journal of the Operational Research Society 51(5): 518-531).

Roysten G., Dost A., Townsend J. and Turner H. (1999) *Using System Dynamics to help develop and implement policies and programmes in Health Care in England,* System Dynamics Review, , Vol. 15, No 3.

Wolstenholme, E. (1990) System Enquiry-A System Dynamics Approach. Chichester, England: John Wiley and Sons.

Wolstenholme, E.F. (1993) *A Case Study in Community Care using Systems Thinking,* Journal of the Operational Research Society, Vol. 44, No. 9, September, pp 925-934.

Wolstenholme, E.F. (1996), *A Management Flight Simulator for Community Care,* In Enhancing Decision Making in the NHS, Ed. S. Cropper, , Open University Press, Milton Keynes.

Wolstenholme, E.F. (1999) *A Patient Flow Perspective of UK Health Service,* System Dynamics Review. Vol. 15, no. 3, 253-273.

Wolstenholme, E.F., (2002) *Patient Flow, Waiting and Managerial Learning - a systems thinking mapping approach,* Working paper, Symmetric SD Ltd..

Wolstenholme E. F., (2003), *Towards the definition and use of a Core Set of Archetypal Structures in System Dynamics,* System Dynamics review Vol. 19, No. 1, Spring: 7-26.

Wolstenholme, E.F. (2005) Monk, D., McKelvie, D. and Smith G., *Coping but not Coping in Health and Social Care – Masking the Reality of Running organizations beyond Design Capacity*, Paper presented to the System Dynamics Conference, Boston, USA.