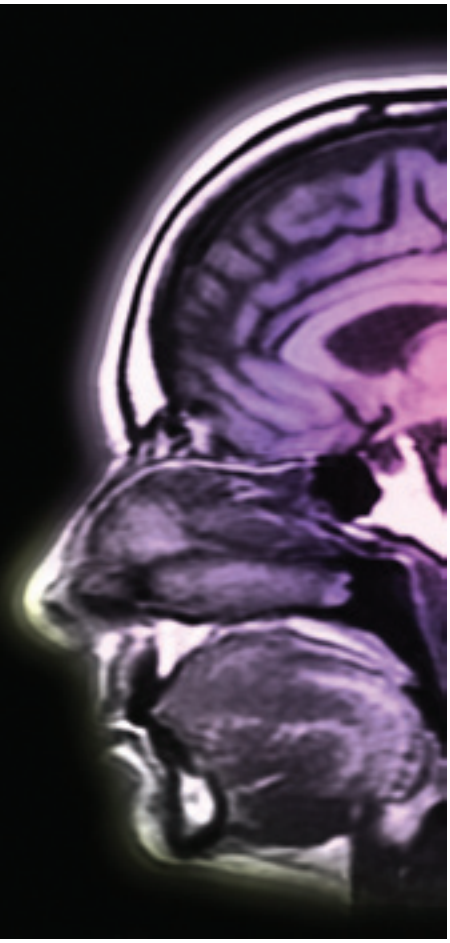
NHS SCOTLAND

The Scottish Government
Health Delivery Directorate
Improvement and Support Team

Understanding Capacity and Demand

A resource pack for healthcare professionals



healthier scotland
SCOTTISH GOVERNMENT

The Scottish Government
Health Delivery Directorate
Improvement and Support Team

Understanding Capacity and Demand

A resource pack for healthcare professionals

ISBN: 978-0-7559-5579-4

The Scottish Government
St Andrew's House
Edinburgh
EH1 3DG

Produced for the Scottish Government by RR Donnelley B53620 12/07

Published by the Scottish Government, December 2007

Further copies are available from
Blackwell's Bookshop
53 South Bridge
Edinburgh
EH1 1YS

The text pages of this document are printed on recycled paper and are 100% recyclable

Contents

A foreword from Richard Steyn

As I have progressed through my medical training from medical student through general practice and subsequently as a surgeon, the emphasis has always been on structured history and examination followed by appropriate investigations and development of a management plan based on knowledge of the appropriate evidence base.

Conversely, throughout my career I have constantly been exposed to the assumption that waits and delays are inevitable and that it is a fact of life that we will effectively ration care using waiting lists. As a cardiothoracic surgical trainee, and during the initial period as a consultant, I shared the cultural belief that if only our funding was top-sliced off the budget and we were left to manage everything ourselves, we would be fine and could manage our waiting lists at a safe level.

Then I attended a workshop on demand and capacity and I was challenged to justify my beliefs. I had no evidence that waits and delays were inevitable. I had no evidence that ring-fencing resources for our own practice was beneficial. We were not applying any scientific method to our thinking around the organisational structure and management of our services and we were certainly not recommending plans developed with an understanding of the evidence base.

As I have investigated further I have been constantly surprised at what can be achieved by careful investigation, assessment and planning of the delivery of clinical services and how "impossible" waiting lists can be reduced or removed. I have been equally surprised how provision of additional resources to the wrong points in the pathways can actually make services worse.

Without a shadow of a doubt the best people to understand and help plan the service are the clinical and managerial teams providing that service and I commend to you to take an evidence-based approach. Achieve the best with the resources you do have, support business development plans with good evidence of need, and instead of being resigned to your work with the frustrations of waits and delays, redesign your work and look forward to coming to work every Monday!

Introduction

This publication is a valuable resource for anyone interested in improvement work in a healthcare setting. It incorporates two DVDs previously published separately by the Diagnostics Collaborative Programme (part of the Improvement and Support Team) but which cover a similar theme, that is, the use of capacity and demand information.

Understanding Capacity and Demand

The first DVD (attached inside the front cover) is a recording of a masterclass lecture given by Mr Richard Steyn, a Cardiothoracic Surgeon, at the first Learning Workshop of the Diagnostics Collaborative Programme in 2006. This gives the viewer an entertaining insight into the practical ways that capacity and demand information can be used to great effect, including how to overcome some of the possible obstacles and pitfalls that are commonly encountered by healthcare staff trying to improve their services. A biography of Mr Steyn is included in this document, along with a detailed summary (with timings) of the points he makes during the 2½-hour lecture. The reader is encouraged to watch this DVD to fully appreciate the comments, models and examples provided by Mr Steyn who is a great believer in the use of information for improvement. Indeed, having initially been a sceptic he went on to build several computer models to demonstrate the theory, and has worked with many hospitals to improve their services. This DVD is bookmarked by chapter to allow the viewer to dip in and out of the wealth of material available. A copy of Mr Steyn's presentation slides is included on the DVD.

Using Demand, Capacity, Activity & Queue (DCAQ): The Experience in Scotland

This second DVD (attached inside the back cover) contains a collection of interviews with staff working with the Diagnostics Collaborative Programme across NHSScotland who have been using demand and capacity information to improve their diagnostic services. The interviewees include service managers, project managers, consultants and waiting list coordinators, each with an enthusiasm for improvement and practical experience of using information for improvement. A summary of the interviews and participant contact details are provided in this document.

The aim of the Improvement and Support Team (IST), part of the Health Delivery Directorate of the Scottish Government, is to share and spread good practice. The subjects covered in this document will never date, as we believe there is always room for improvement, and information is always needed to achieve this aim. Please feel free to distribute this resource as a useful reference and teaching tool. For more information about any of the programmes supported by the IST please contact istmailbox@scotland.gsi.gov.uk.

Biography of Richard Steyn MS FRCSEd(C-Th) FIMCRCSEd MRCGP DRCOG

Richard S. Steyn qualified as a doctor from Aberdeen University in 1984. After completing vocational training in rural general practice (Taynuilt Medical Practice) he returned to surgical training in Aberdeen, Liverpool, Manchester and Birmingham. In August 1999, he was appointed Consultant Thoracic Surgeon at Heart of England NHS Foundation Trust. His main clinical interests are oesophageal cancer and cardiothoracic trauma. He is also an active pre-hospital doctor with West Midlands Ambulance Service and Warwickshire Ambulance Service and was awarded the Queen's Golden Jubilee Medal for his services.

He has been involved in the English cancer services collaborative from Phase I as a tumour specific lead (lung), and in Phase II as a regional clinical lead. He has led several initiatives including patient-led follow-up and taping patient consultations. His main interest however has been researching healthcare processes with particular emphasis on understanding the effects of variation and flow on the management of demand and capacity and the resulting waiting lists. He has developed several computer models to facilitate the understanding and teaching of the concepts that must be addressed to allow effective use of healthcare resources. He was appointed National Clinical Lead Demand, Capacity & Patient Flow with the NHS Modernisation Agency in December 2001 and National Clinical Lead Cancer Modernisation in March 2003. More recently he has been appointed as an Honorary Senior Lecturer with the University of Warwick.

Increasingly he is lecturing and consulting on healthcare process design and demonstrating these models both within the UK and internationally to facilitate the understanding of demand and capacity management and planning, queuing theory, theory of constraints and lean thinking as it relates to healthcare.

Richard's models and other presentation materials can be found at his website: www.steyn.org.uk.

Understanding Capacity and Demand – a masterclass lecture

Learning points

The following pages give a summary of the learning points made in each of the chapters on the DVD with the start time and length of the chapter included. This gives the viewer an idea of the structure of the lecture so that it can be watched in full or in part; it also enables the DVD to be used as a teaching resource. There are two question and answer sessions within the lecture. The questions posed by the audience are marked as Q1 to Q8.

Using the 'View Chapters' selection on the main DVD menu you can choose which sections of the lecture to watch. Alternatively select 'Play Lecture' to watch the entire masterclass, or select 'Presentation' and follow the instructions provided to view the slides Richard used for his lecture.

Chapter	time	length
What's the problem?	**00:00:00**	**00:01:16**
Is the demand on your system greater than the capacity? Is the link between demand and capacity really understood?		
Running a healthcare system	**00:01:16**	**00:04:18**
Is it possible to run a healthcare process with no wait and no waste? Using a model of how patients flow through a 5-step pathway, the audience is asked to run the model so that the end result is no waste for the service and no wait for patients.		
What's the result?	**00:05:34**	**00:03:48**
The health service often works inefficiently because we don't properly understand how it actually works. What information are you using to try and understand the system? Are you basing your decisions on useful information or meaningless data?		
Hitting targets	**00:09:22**	**00:00:52**
Distorting the system to make sure targets are met is often the result of not understanding the system in the first place.		
Consequences of not understanding variations	**00:10:14**	**00:00:46**
What happens because we don't understand our systems? Are you looking for trends where none exist, using past activity data to manage your service? How often have you had business cases rejected due to a lack of quality information?		
Interpreting data	**00:11:00**	**00:01:50**
What data do you need to properly understand your systems and how should you display and interpret it? Are you able to show that a change you have made has definitely made an improvement?		
Statistical process control (SPC)	**00:12:50**	**00:02:33**
What is statistical process control (SPC) and why is it a useful way to analyse data? Monitoring data over time gives you a better picture of the variation in your service.		

Understanding variation	00:15:23	00:01:44

What is variation and how does it affect your service? Learn about common and special variation and how you should interpret variation to make the best use of your resources.

How did the NHS get into this mess?	00:17:07	00:02:41

Summarises the discussion so far. How often do silo thinking and targets create pressure and problems? Making sure you understand the system helps you to quantify what really needs to be done.

Demand and capacity	00:19:48	00:02:05

Demand, capacity, activity and queue are defined here, with an explanation of how to measure them. Do you really know the demand for your service? Should you be counting patients or time? What are the constraints on your capacity?

Why do queues form?	00:21:53	00:04:42

Do queues only form when demand is greater than capacity? Here is a demonstration of how queues form due to the mismatch in the variation in demand and the variation in capacity. Are queues really necessary? Are you relying on waiting list initiative work to keep the queue under control? How predictable is your service?

Utilisation	00:26:35	00:05:10

Is working at 100% utilisation efficient? Are your staff working harder with no obvious improvement to the service? This chapter discusses the impact that pressure to utilise resources fully has on the system. What's wrong with batching? What are the costs of having a queue?

How do we traditionally respond?	00:31:45	00:12:28

How does the health service traditionally respond to having a waiting list? Here are some illustrations of typical responses including delaying the patient, forced booking, carving out capacity, using waiting list initiatives and pressurising the system.

The road to ruin	00:44:13	00:01:25

Failing to understand the system leads to a vicious circle of increased variation and falling activity resulting in longer waits and cutbacks. Does this sound familiar?

So what should we do instead?	00:45:38	00:16:22

Q&A with the audience about alternative ways of managing and measuring systems in order to provide quality services. Topics covered include: investing in people; reducing demand; reducing variation; predicting demand; process design.

What should we do instead? (summary of the above chapter)	01:00:30	00:01:30

Top tips include: focussing on the quality; managing the bottleneck; aiming for effective outcomes.

The patients and the process view	**01:02:00**	**00:15:40**

How does the patient view your process? Have you considered following the pathway as a patient? By looking at the number of steps in your process you can identify how many of them actually add value for the patient rather than being the result of a work-around and ultimately wasteful. This chapter looks at the probability of performing each step in the process successfully. How often do you get the right result first time? Reducing the number of steps will increase the chances enormously. Watch out for special cause variation but aim to reduce the common cause variation which affects the majority of patients.

Effects of pooling	**01:17:40**	**00:11:53**

A computer model demonstrates the effect of pooling on a queue. Do you have several queues for the same service? Reducing the number of queues will reduce the length of wait – compare your queuing system to that used by the Post Office. Have you looked at the profile of your service? What procedures are the most common?

What do patients think?	**01:29:33**	**00:06:41**

This chapter discusses how patients feel about waiting, pooling of lists and travelling for treatment. Here are some illustrations of the effects of seeing patients in turn and matching capacity and demand. How often do you increase demand based on prior knowledge – for example, Breast Awareness week?

Emergency and elective admissions	**01:36:14**	**00:02:13**

Which do you believe to be more predictable – emergency or elective admissions? The answer may surprise you.

See "today's" demand "today"	**01:38:27**	**00:01:22**

Stop the queue getting out of hand by seeing today's demand today. How should you work out what capacity is needed? By looking at the fluctuation in demand you can calculate the capacity required to maintain a certain level of service. Remember no wait = waste, but no waste = wait.

Setting the right capacity	**01:39:49**	**00:07:30**

This computer model demonstrates the effect of capacity on a queue. How should you calculate the required capacity for a service? Here is an example of setting the capacity for a CT service using a specific equation (the 80% rule, see equation below) to take into account the variation in demand.

theoretical capacity = minimum demand + (maximum demand – minimum demand) * 0.8

Behaviour change	**01:47:19**	**00:01:40**

Have you noticed the number of hospital cancellations rising and activity falling when the pressure is on to reduce waits and maximise available capacity? This chapter discusses the use of available capacity and the effect it has on staff behaviour.

Length of stay	**01:48:59**	**00:06:40**

How does the way we run our system affect the patients' length of stay? Here are some ideas for reducing lengths of stay including improved discharge planning, admission on the day, increased ward rounds improving discharge patterns.

The system approach	**01:55:39**	**00:00:16**

Using the system approach improves time, cost and quality whereas the focus is usually on either cost or time.

QUESTION & ANSWER SESSION

Q1: How much data do you need to have to know you fully understand your variables in your system?	**01:55:55**	**00:02:38**
Q2: Is the 'no queue' principle a sensible thing?	**01:58:33**	**00:05:34**
Q3: Is protected time for radiologist reporting helpful and how does it fit into the model?	**02:04:07**	**00:03:48**
Understanding the system	**02:07:55**	**00:07:35**

This chapter summarises what to do in the short and long term to improve access to services.

Q4: If your unit has a system that is currently working well, do you still continue to collect data?	**02:15:30**	**00:02:17**
Q5: How, in practice, can you go about tackling considerable variability?	**02:17:47**	**00:01:36**
Q6: How do we influence the approach to tackling waiting lists?	**02:19:23**	**00:02:53**
Q7: Should we wait to collect capacity and demand data until after the summer months when activity is different because there are lots of staff away?	**02:22:16**	**00:01:56**
Q8: Can you give us simple tips on how to measure the effect holidays have on the capacity/demand equation?	**02:24:12**	**00:03:33**

Understanding Capacity and Demand – a masterclass lecture

Detailed points

For a more in-depth guide to the points made by Richard in each of the chapters you should refer to the following pages.

Using the 'View Chapters' selection on the main DVD menu you can choose which sections of the lecture to watch.

Chapter	time	length
What's the problem?	**00:00:00**	**00:01:16**

- "demand is greater than capacity"
- a lack of understanding of demand

Running a healthcare system	**00:01:16**	**00:04:18**

- modelling how patients flow through a 5-step pathway
- models available from www.steyn.org.uk
- the audience attempts to run the model with the aim of **no wait and no waste**
- where is the bottleneck?
- resources are put into the front-end and all perceived waste in the system is removed

What's the result?	**00:05:34**	**00:03:48**

- chaos and increased cost
- delays and deteriorating patient conditions
- reduced staff morale, frustration
- over-complicated board reports
- lack of understanding
- incorrect assumptions and decisions made
- emphasis on targets
- no change to processes
- examples of unintelligible analysis

Hitting targets	**00:09:22**	**00:00:52**

- to improve performance against targets we distort the system and distort the data, usually without understanding the system in the first place
- examples of methods that have been used to meet the 4-hour A&E target:
  - remove the wheels from the trolleys!
  - relabel corridor space as ward!
  - don't accept the patient from the ambulance!

Consequences of not understanding variations	**00:10:14**	**00:00:46**

- non-existent trends are found
- blame or credit non-responsible individuals
- barriers go up, morale goes down, create fear
- cannot use past performance, cannot predict the future, cannot significantly improve the system
- business cases are based on flawed data and presented badly

- resources are consequently not made available
- using the right information and presenting it clearly in a business case will increase the likelihood of success

Interpreting data	**00:11:00**	**00:01:50**

- pitfalls of poor data presentation
- statistics for comparison
- is a change really an improvement?
- example:
  - average wait reduced from 70 days to 35 days
  - demonstrated with different charts
  - importance of monitoring over time
  - stop monitoring, effect is reversed

Statistical process control (SPC)	**00:12:50**	**00:02:33**

- an explanation of statistical process control and why it is a useful way to analyse data
- plot data over time to examine the variation
- examples:
  - recording the temperature of a patient
  - a trip to work

Understanding variation	**00:15:23**	**00:01:44**

- every process displays variation
- common cause – consistent pattern, by chance
  - should not make decisions based on slight changes
- special cause – assignable, pattern changes over time
  - if it has a positive effect incorporate it, if not, remove/avoid it
- example: the pumpkins
- SPC rules to detect whether variation is common or special cause
- SPC points out the questions to be answered, not the answers themselves

How did the NHS get into this mess?	**00:17:07**	**00:02:41**

- a lack of understanding of the system
- a lack of real information despite lots of data
- huge pressures
- silo thinking – in the model, step 2 cannot solve the problem by themselves
- using inappropriate measures of performance
- with a "hit the target" mindset:
  - waiting lists are due to a lack of resources
  - so we increase resources
  - which demands value for money
  - we think 100% utilisation proves maximum efficiency
  - the financial deficit increases whilst patients continue to wait
- SPC example:
  - urgent GP referrals to colorectal cancer surgery – quantify the real problem if all patients to be seen within 62 days

Demand and capacity	**00:19:48**	**00:02:05**

- what really is the demand for your service?
  - usually quoted as average activity for the past 3-6 months
  - usually wrong
- demand = what we should do
- capacity = what we could do
- activity = what we did do
- queue = what we should have done
- measure each in the same time frame to compare
  - example: compare 5 oesophagoscopies with 5 oesophagectomies
- constraint on capacity = kit or staff
- commonest constraint = staff
- commonest request for additional resource = kit
  - but kit is sometimes handed back, as staff was the resource required

Why do queues form?	**00:21:53**	**00:04:42**

- demand is greater than capacity (sometimes)
- normally due to the **mismatch** between the variation in demand and the variation in capacity
- having a queue signals high utilisation, it keeps us busy
- the service was designed to have a queue
- demonstration of the queuing model
  - explanation of waiting versus waste
  - if demand is truly greater than capacity, the queue rises constantly
- "waiting waits, waste is gone" i.e. demand pushes forward but capacity is lost
- effects of waiting list initiatives:
  - impact downstream – no extra funding, queue forms
  - end of initiative – holidays, reduced efficiency
  - higher impact on normal running of the department
- examples:
  - breast clinic – demand is predictable, clinic slots are affected by holidays
  - hospital admissions and discharges – less variation in admissions
  - elective and emergency admissions – emergency admissions are more predictable than electives (note Richard misquotes here but reiterates correctly)

Utilisation	**00:26:35**	**00:05:10**

- is a very judgemental measure
- the assumption is that efficiency = 100% utilisation
- and that having a queue means the resource is fully utilised
- this leads to a push for more resources or a redistribution of resources
- which keeps the pressure on to have a queue
- example:
  - surgery lists expand to fill time available, but cases per list will fall
- leads to batch logic
  - variable demand placed into set clinics
  - which leads to surges in demand
  - example:
    - once a month rheumatology clinic increases demand on radiology

- is NOT efficient; low unit cost BUT patients wait
- there are associated costs
  - of managing the queue
  - of deterioration in the queue
  - of impact on downstream capacity

- examples:
  - CT radiologist spending 2-3 hours per week managing requests
  - SPC: CT exam to reporting time

- can lead to silo thinking and money wasted
- utilisation is the driver to pressurise the system – do we really want 95-100% utilisation?
- example:
  - compare with jet fuel to Australia

How do we traditionally respond?	**00:31:45**	**00:12:28**

- we delay the patient
  - keep minor injury patients until there are enough to warrant a medical staff member to treat them

- we use forced booking
  - book extra patients into a clinic; everyone waits longer, frustration increases, leads to blocking tactics

- we carve out capacity
  - model shows the effect of making an 'urgent' queue; routines wait even longer; add an 'urgent urgent' queue; queue goes out of control
  - which leads to increased demand
    - lung cancer slots in radiology
    - orthopaedic referrals
  - results in waste
    - orthopaedic patients no longer fit leading to cancelled surgery
  - results in churn
    - patients still waiting fall ill
    - patients start phoning in to be appointed sooner
  - patients start to appear the same
    - cancer patients – urgents and routines wait the same time
  - carve out affects not just lists
    - CT scanning schedule – 73 queues!
  - dangers of carve out

- we use waiting list initiatives
  - perverse incentives, not cost effective, impact downstream

- we pressurise the system
  - performance manage with targets – "hit the target, miss the point"
  - leads to bullying, reduced quality, lower staff morale, blocking tactics, increased costs
  - example: two emergency departments and their 4-hour trolley waits

The road to ruin	**00:44:13**	**00:01:25**

- the failure to understand the system leads to increased variation within it – creating a vicious circle
- where capacity plans and contracts are based on average past activity
- leading to the failure to account for the variation in demand and capacity
- which means we fail to deliver the required activity
- so our income is lower than expected
- and our waiting times are longer than guaranteed targets
- leading to increased overtime and initiative work
- and more waiting list validation and breach analysis
- the costs go up
- so cost-cutting begins
- with risk of redundancies and cuts to services
- which reduces the capacity
- and increases the variation in capacity!

So what should we do instead?	**00:45:38**	**00:16:22**

This is a question & answer session with the audience about alternative ways of managing and measuring systems to provide quality services:

- "invest in people":
  - invest in the right place and remove waste elsewhere
  - change where your money is if it's not working
  - usually require skill not kit
- "reduce demand":
  - it's not the patient's fault they are waiting
  - does the length of wait really matter?
  - "right department, right time"
  - earlier treatment, easier to treat, shorter length of stay
  - self-monitoring, early warning systems, triggers treatment
- "reduce the variation within our control":
  - examples: Christmas Eve discharges; ward rounds affecting discharge patterns; weekend ward rounds
  - use discharge planning to reduce length of stay rather than discharging patients when needing the bed
- "enhance primary and secondary care relationships"
  - maximise the utilisation of primary/secondary care capacity
- "predict demand"
  - example: winter pressure planning
  - watch out for increased demand due to lower waits, is it for the right reason?
    - examples: 64-slice CT; lower threshold for cataract surgery; increased risk of inappropriate surgery and death
  - smooth demand
    - example: stagger GP home visits which were creating a surge in demand at A&E
- "design the process with the strategic objective in mind"
  - compare the design and production of a luxury car with the bolt-on process in healthcare

- "know your strategic objective and how you relate to it"
  - otherwise all staff try to achieve their own aim and pull in different directions
  - examples: financial versus clinical service; an activity-based service moving to a demand-responsive service

What should we do instead? (summary of the above chapter)	**01:00:30**	**00:01:30**

- focus on quality
- manage the bottleneck
- plan for no queue
- focus on flow
- know your service to manage and plan
- aim for an effective outcome

The patients and the process view	**01:02:00**	**00:15:40**

- highly recommend following the pathway as a patient
  - example: Chairman acting as a lung cancer patient
- aiming for 18 weeks and 62 day pathways
- look at the system as steps in a process
  - example: urology clinic – 109 steps!
- get outside opinion
  - example: signage, directions
- focus on the value and remove the waste in the process
- are the solutions what the patient would want?
  - example: Radio 1 in HDU
- what is the probability of performing each step in the process successfully?
  - example: if we want 9 out of 10 patients to be treated perfectly through a 100 step process, the error rate of each step needs to be less than 1 in 1000
  - compare with the electronics or aircraft industry with 3.4 defects per million
  - we are currently lucky to have 1 in 10 patients without any error
- increase the probability of success by reducing the number of steps before chasing the quality
  - examples: patients take their own notes to the appointment; one-stop clinics; lung cancer pathways; direct referral to rapid-access clinic
- investigate special causes but also manage the normal variation down
  - don't assume the answer lies within the department
  - example:
    - time to surgery for colorectal cancer, wanted extra staff, theatre time, beds
    - BUT DCAQ showed endoscopy activity was greater than demand, and yet there was still a queue
    - so surely they must need more endoscopists!
    - BUT activity was less than actual capacity, which was less than the theoretical capacity
    - BECAUSE of carve out
    - with 73 queues the service was impossible to manage!
    - THEREFORE there was no need for extra theatre staff etc.
  - example:
    - long queues, must need more endoscopists
    - actually only 2 toilets in endoscopy unit
    - answer – get patients to use bowel preparation at home!

Effects of pooling	**01:17:40**	**00:11:53**

- model demonstrates the effect of pooling on a queue
  - 8 independent clinics with the same set-up, aiming for a 2-week wait but all go over the 2-week wait
  - if all the referrals are pooled the process never goes over a 2-week wait, with the same amount of waste
- reduce carve out
  - example: ovarian cancer referrals, pooling reduced waits from 130 days to 30 days
- share diaries to offer patients the first available dates
- discussion:
  - patients are educated to prefer certain clinicians by their own doctors
  - need to trust your colleagues
  - you should not tolerate poor practice
- pareto analysis:
  - 20% of procedures account for 80% of the work; 7% for 50%;
  - consider whether the rare procedures should be performed at all
  - should they be done in one hospital?
  - pool the routine procedures
- know what you know – transfer the patient at the appropriate time
- "no service should depend on one individual"
  - spread the skills across the department(s) to create a robust and sustainable service
- patients can choose to wait
- "don't just take the solutions and apply them to your service, take the thinking and generate your own solutions"

What do patients think?	**01:29:33**	**00:06:41**

- how do patients feel about waiting, pooling of lists and travelling for treatment?
- don't make assumptions about patients – ask them!
- for routine tests, I want to be close to home; for specialist tests, I will travel
- patients will travel some distance for pre-operative assessment, using it as a test run
- patients don't want surprises, they will normally accept being treated by your colleague if necessary if informed in advance
- some patients with dependents need to plan in advance
- would probably not wish to undergo a major procedure 2 days before Christmas
- taking patients in turn leads to a dramatic reduction in waits and smoothes the variation in waits
  - example: Scarborough Hospital
- matching variations in demand and capacity can have a big impact
  - example: breast clinic – moved from 2 clinics per week with 54 slots to 3 clinics per week with 48 slots; waits reduced from over 2 weeks to 5 days, no vetting required
- use known indicators to plan for increases in demand
  - examples: Breast Awareness week; calls to NHS 24; changes in air temperature and pollution
  - compare with a supermarket on a hot and sunny local derby day

Emergency and elective admissions	**01:36:14**	**00:02:13**

- emergency admissions are more predictable than elective admissions
  - examples:
    - a North Warwickshire hospital admissions on 14th January
    - Manchester Royal Infirmary had "nightmare" Mondays with too many elective admissions
- reducing the variation in elective admissions can reduce the expected number of beds required (but keep the slack)
  - example: by reducing the variation in the elective admissions the expected number of beds required fell from 78 to 68 per day

See "today's" demand "today"	**01:38:27**	**00:01:22**

- where "today" could mean this hour, this week or this month depending on the type of service you provide
- never let the queue get out of hand
- set your capacity at 80% of the variation in demand and your queue will stay under control
  - example:
    - if demand varies between 95-105, capacity should be set at 103
    - if demand varies between 50-150, capacity should be set at 130
    - even though both services have an average demand of 100

Setting the right capacity	**01:39:49**	**00:07:30**

- modelling the effect of capacity on the queue
- set your required capacity by calculating 80% of the variation in demand
  - **= minimum demand + (maximum demand – minimum demand)*0.8**
  - but is a wait acceptable?
  - capacity should depend on the service – for resuscitation you would want 100%
- reduce the impact of holidays using cross-cover, increased capacity beforehand, locum staff
- manage annual leave – don't allow the majority of staff to go on leave at the same time!
- have the right person doing the right job at the right time
- use the right data (not just numbers of patients, but time required to treat the patients)
  - example: CT demand, average 79 requests per week, how much capacity is required?
    - demand = requests x scan time
    - requires 2 extra hours a day – is this cheaper than having to prioritise lists and perform waiting list initiatives?
- know what level of wait you are prepared to accept – balance wait versus waste
  - no wait = waste
  - no waste = wait

Behaviour change	**01:47:19**	**00:01:40**

- discussion of the use of available capacity and the effect this has on staff behaviour
- non-linear relationship between cancellations and occupancy
  - once occupancy goes over 85% behaviours change
  - the onus becomes one of bed-blocking:
    - bring patients in early, keep them in longer, don't accept A&E patients
    - no co-operation

Length of stay	**01:48:59**	**00:06:40**

- how does the way systems are run affect length of stay?
- examples:
  - average length of stay by day of admission – similar patients are treated differently
  - reducing the length of stay for a large cohort of patients may increase your average length of stay
    - improvements send out the wrong message if interpreted incorrectly
  - don't admit if you don't have to – bring the patients back to a clinic in the morning
  - admit on the day of surgery not the night before
  - aim to reduce the average length of stay of 80% of patients (exclude the long-stayers with social or medical problems) – costs saved can be redistributed elsewhere
  - cost analyses of length of stay for
    - hip replacement
    - mental health bed occupancy
  - improve discharge planning but check there is no increase in readmission rates

The system approach	**01:55:39**	**00:00:16**

- the system approach improves time, cost and quality whereas the focus is usually on either cost or time

QUESTION & ANSWER SESSION

Q1: How much data do you need to have to know you fully understand your variables in your system?	**01:55:55**	**00:02:38**

- use "crude measures of the right thing rather than precise measures of the wrong thing"
- use the data or the quality will remain appalling
  - the most common procedure coded in hospitals is 'not otherwise specified'
  - ensure that everyone is aware of the need for good quality data and feed it back to them
- statistical process control rules suggest having at least 20 points to detect trends
- use the correct time frame
- find an analyst – maybe this has already been done!

Q2: Is the 'no queue' principle a sensible thing?	**01:58:33**	**00:05:34**

- if you remove the waiting list, there will be a surge in demand due to unmet demand
- once this surge passes the level of demand will reset – it may be slightly higher or lower
  - example: CT – access times fell so requests dropped – if the test is required it will be available rapidly; unnecessary requests disappeared
- you may need a lead time to organise the patient (e.g. bowel preparation)
- but why shouldn't a lung cancer patient be seen today?
- on the other hand, the patient may choose to wait (compare GP appointments)
- why not use full booking?
- why worry about loss of control? there's always demand out there!
- no waiting = no worry about taking leave at short notice
- more slack time = better quality service, better motivation, more time for training
- move to day case treatment, even thoracic surgery!
- prove you need a queue!

Q3: Is protected time for radiologist reporting helpful and how does it fit into the model?	**02:04:07**	**00:03:48**

This is a wider discussion of the difference between carve out and segmentation.

- what is the benefit? does it benefit one group of patients or all patients?
  - examples: minor injuries "see and treat"; cancer patients
- segmentation:
  - reduces the variation in the demand and the process time
    - example: uninterrupted reporting allows for more reports to be completed than if the radiologist is interrupted every 30 seconds
  - improves the flow for all patients
  - increases the return on net assets
- carve out:
  - does not control variation, may make it worse
  - the flow for one group of patients is improved to the detriment of another group of patients
  - capacity is wasted
- to decide if it's right you must monitor the impact in your department
- example: in Heartlands Hospital, simultaneously have uninterrupted reporting and other staff available for questions and support

Understanding the system	**02:07:55**	**00:07:35**

- the root cause of delays is variability and high utilisation, not volume
- in the short term, aim to optimise current capacity
  - reduce the number of steps and queues
  - treat in turn, plan discharges
  - maximise skill use, pool capacity
  - get it right first time
- in the long term, plan for no queue (or minimal queue)
  - measure and shape demand
  - plan capacity
  - reduce variation in demand and capacity
- back to the healthcare computer model – where is the bottleneck?
  - step 5 (discharge) were working with no waste
  - reduce the impact of holidays
  - reduce the impact of variation
  - "balance" the line
  - waste will reduce, wait will reduce to a lower level
  - gradually feed the backlog into the system
- the demonstration uses one flow whereas radiology, endoscopy and pathology are "hubs" in the hospital with more than one flow which means they are often the centre of attention when actually the problems are caused elsewhere
- needs a global view
- "chain of dependence"

Q4: If your unit has a system that is currently working well, do you still continue to collect data?	**02:15:30**	**00:02:17**

- don't measure everything all of the time
- measure what's relevant
- measure before you begin and as you change
- example: waiting times

Q5: How, in practice, can you go about tackling considerable variability?	**02:17:47**	**00:01:36**

- don't do waiting list initiatives
- if they must be done, give the initiative to the staff with the shortest waiting list to remove the perverse incentive
- also, fund the support services and staff
- focus attention on getting it done right
- focus attention on giving the highest level of service

Q6: How do we influence the approach to tackling waiting lists?	**02:19:23**	**00:02:53**

- challenge senior management
- engage colleagues
- compare with the state of the hospital (take photographs, raise the issue)
- who is providing the advice and guidance to politicians?

Q7: Should we wait to collect capacity and demand data until after the summer months when activity is different because there are lots of staff away?	**02:22:16**	**00:01:56**

- start looking, start measuring, start acting!
- sometimes only need small changes, so why wait?
- aim for quick wins
- get the momentum going
- aim to embed a culture of change and service improvement
- use working hours to meet – will improve productivity – make it part of normal work so staff feel engaged and attend

Q8: Can you give us simple tips on how to measure the effect holidays have on the capacity/demand equation?	**02:24:12**	**00:03:33**

- no easy solution
- watch out for the wandering bottleneck
- recognise where something isn't working and change it
- or, get the service delivering a reasonable level of work
- don't aim for the ultimate service, aim for simplicity
- want the right resource at the right time and right place
  - example: why use nurses to clean theatres, why not have a team of cleaners moving from theatre to theatre at changeover?
- plan ahead – how often is one key member of staff on holiday and nobody knew?

Using Demand, Capacity, Activity & Queue (DCAQ): The Experience in Scotland

Choose 'Select Interview' to jump to a particular section of the DVD or select 'Play All' to watch the entire DVD.

Interviewee	**Précis**
Ian Penman (Consultant Gastroenterologist, NHS Lothian) ian.penman@luht.scot.nhs.uk 0131 537 2477	A consultant's perspective: how meaningful information can expose the good, and the extent of poor practice, in a service; how it can be used to initiate dialogue with staff and to direct sustainable change.
Eleanor Kinghorn and Theresa Holliman (Waiting List Coordinator and Personal Assistant/Waiting List Coordinator, NHS Ayrshire & Arran) eleanor.kinghorn@aaaht.scot.nhs.uk 01292 617067 theresa.holliman@aaaht.scot.nhs.uk 01563 577349	Moving from 'data' to 'information': how communication has led to the better use of clinical and administrative time, improved room utilisation and a more informed service.
Lindsay Potts (Consultant Gastroenterologist, NHS Highland) lindsay.potts@haht.scot.nhs.uk 01463 704000 bleep no. 7064	The inconvenient truth: challenging historical assumptions and behaviour with accurate information to instigate a more streamlined service where both patients and staff benefit.
Anne Haythorne (Charge Nurse/Endoscopy Decontamination Manager, NHS Fife) anne.haythorne@faht.scot.nhs.uk 01383 623623 x 7949	'Simple' changes can make an enormous difference: if the data you receive isn't informing your service focus on the problem area, collect the right information, and feed it back to staff in order to generate a culture of improvement.
Jonathan Procter (Director of Patient Access/ Associate Finance Director, NHS Forth Valley) jonathan.procter@nhs.net 07747 767309	Achieving the biggest 'bang for your buck': how DCAQ is integral to understanding where to improve the use of clinical resources, and how embedding it's collection into normal working practice will ensure the effective monitoring of your service.
Lindsay Campbell (Project Manager, NHS Ayrshire & Arran) l.campbell@aaaht.scot.nhs.uk 01563 825097	Working with the service to make real and sustainable improvements: how a tailored approach will encourage staff to take ownership and draw reasoned conclusions from the information; how staff then feel more confident making suggestions to improve the service.

Interviewee	Précis
Aileen MacLennan (General Manager, Imaging & Clinical Physics, NHS Greater Glasgow & Clyde) aileen.maclennan@northglasgow.scot.nhs.uk 0141 211 4610	Effective information management in a large improvement project: how a good communication structure generates stakeholder involvement, ownership and empowerment within the decision-making process.
Hakim Ben-Younes (Consultant Surgeon/Specialty Clinical Director, NHS Lanarkshire) hakim.benyounes@lanarkshire.scot.nhs.uk 01698 366135 (secretary)	Improving the patient pathway: how accurate information has provoked changes in working practice leading to improved quality of patient care and access to services.
Marie Martin (General Manager, Diagnostics, NHS Greater Glasgow & Clyde) marie.martin@rah.scot.nhs.uk 0141 314 6635	Justifying the case for extra capacity: how the monitoring of demand, capacity and utilisation of existing equipment confirmed the anecdotal lack of capacity, and informed a business case to secure funds to promote a sustainable service.
Paul Duffy (Clinical Director, NHS Greater Glasgow & Clyde) paul.duffy@sgh.scot.nhs.uk 0141 201 1558	Using a whole systems approach: the importance of having objective evidence to challenge inefficiencies within the system, and to establish a planned and flexible service.

An A-Z of Demand and Capacity terminology

Term	Meaning
80% rule	By monitoring the total demand on a service over a period of time it is possible to calculate the theoretical capacity that would give the process the flexibility to cope with the variation in demand and maintain a manageable queue. **theoretical capacity = minimum demand + [(maximum demand – minimum demand) * 80/100]** This figure is a guide; if you need no wait then you should be prepared to deal with 100% of your demand; if you are prepared to have a slightly longer wait for the service then this figure could be reduced.
Activity	The work done; the patients seen through the process; the actual capacity that is used to see patients. Aim to reduce lost activity due to patient DNAs and cancellations. To measure activity multiply the number of patients seen by the time it took to process them.
Bottleneck	A bottleneck is where queues form; they slow down the whole process. There are two types of bottleneck: the *process* bottleneck which limits the rate of the process, i.e. the step that takes a significant amount of time; and the *functional* bottleneck which occurs when a resource is used by more than one process, e.g. porters. It is often relatively cheap or free to remove or reduce the effect of a bottleneck.
Capacity	All the resources required to do the work; including rooms, staff and equipment; *theoretical* capacity takes into account the availability of rooms, downtime for maintenance of equipment, staff holidays and study leave etc. *Theoretical* capacity is the maximum capacity that could be utilised if lists were run to time, fully booked, equipment was available etc. *Actual* capacity may be lower due to staff unavailability, equipment breakdown, adverse incidents etc. To calculate theoretical capacity multiply the number of rooms and pieces of equipment by the time available to the people with the necessary skills to use them, e.g. a theatre may be physically available 24 hours a day, but if it is only staffed for 7 hours a day, 5 days a week for 42 weeks then this would allow a theoretical capacity of 1470 hours per year. Theoretical capacity should ideally be calculated using your demand (see the 80% rule above).

Carve out	Capacity that is retained for a particular type of patient or test (e.g. urgent slots on a clinic template, slots set aside for knee operations). Carve out causes queues to lengthen as slots go unused or are misused (patients jumping the queue) which leads to increased variation. The flow of one group of patients is improved at one bottleneck to the detriment of others.
Churn	Churn starts when patients are waiting a long time and are not called in date order. More and more referrals may be flagged as urgent, and even those who are flagged as routine may jump to the top of the queue by phoning the secretary for an earlier appointment. Patients who don't chase the system fall further and further down the list, to the point that their health may deteriorate, resulting in an emergency admission, or they may get treated privately.
Constraint	Ultimately restricts the capacity of the system, e.g. the number of theatres, the number of surgeons. These are not easily increased, so you should aim to utilise them effectively.
DCAQ chart	A graph plotting the demand, capacity (theoretical and actual), activity and queue on a single graph in the same unit of measurement – they are all inter-related. If activity is greater than demand for a sustained period, you should see the queue fall. If capacity falls (for example over holiday periods) you can compare it to the demand which may not have such a significant reduction. This will lead to an increased queue. It is also worth measuring the number of DNAs and cancellations as these will have an impact if not managed.
Demand	All the requests into the service from all sources. Remember the repeat patients who need to be seen albeit at a specified time. Use demand to set the theoretical capacity (see the 80% rule above). To measure demand multiply the number of patients waiting by the time it will take to process them.
Flow	Improve the flow of patients along a pathway by reducing or removing bottlenecks, reducing hand-offs, removing non-value-added steps, reducing variation, managing capacity and demand, and increasing the probability of getting it right first time.
Mismatch	The mismatch between the variation in demand and the variation in capacity will usually result in a queue.
Non-value-added steps	Steps within a patient pathway or other direct process that could be deemed as being time-wasting and of no benefit to the patient or process, e.g. 6 steps to appoint a patient, 5 steps to vet a request. Aim to eliminate the non value-added steps to streamline the process and shorten the time taken.

Pareto analysis (similar to the Glenday Sieve analysis)	By looking at your service you will find that only 20% of procedures contribute to 80% of the workload; 7% contribute to 50%. Consider whether your service should be performing the very rare procedures. Concentrate on improving the flow for the high-volume procedures – this is where you will have the biggest impact on waiting times.
Pooling	Rather than each consultant maintaining a separate queue, consider pooling all straightforward referrals into one queue based on urgency and date of receipt. By pooling referrals the queue will stabilise and patients will wait equitably. This in turn will reduce churn and may additionally remove the need for an urgent category.
Process-mapping	By looking at the process from the patient perspective as a series of steps it is possible to identify what steps add value to the patient pathway and what simply add time and bureaucracy into the process. Remember to consult all services, staff and other stakeholders who may be affected by changes you make.
Queue	The demand which has not yet been dealt with (also known as backlog), i.e. the patients still waiting to be seen. Queues form whenever demand exceeds capacity, or, more typically, when demand and capacity are mismatched. To measure the queue multiply the number of patients still waiting by the time it will take to process them.
Run chart	A line graph plotting the item you are interested in measuring as consecutive points (e.g. time from referral to treatment; number of patients waiting). Helps to identify common and special cause variation.
Segmentation	Segmentation separates the process of care along the whole pathway for one group of patients in order to improve the overall flow of patients but not at the expense of other patients.
Silo thinking	Aiming to improve the processes that affect you rather than aiming to improve the entire patient pathway or process; failing to recognise that decisions and changes you make can affect other departments and services outwith your control.
Statistical Process Control (SPC)	A more sophisticated form of a run chart which measures the limits and the average of the process; has set rules to identify trends, for example, when a process is out of control (or a change has had an effect), and gives more information about the work that may be required to bring a process into more acceptable limits, i.e. to reduce the common cause variation. It can also provide evidence that changes made to a process have actually improved it.
Treat in turn (first in, first out)	Where waiting times are long the temptation is to create urgent queues. However, all research has shown that this simply leads to longer waits for the non-urgent patients. Treating patients in turn will reduce the churn and reduce the waiting time for all patients in the longer term.

Utilisation	A measure of how much capacity is used; 100% is often seen as the ideal, but actually the process should allow for some flexibility as the pressure to fully utilise resources will lower staff morale and incite behavioural change. To calculate utilisation divide actual activity into actual capacity and multiply by 100 e.g. 480 minutes of activity out of 540 staffed minutes gives 480/540 * 100 = 89% utilisation.
Variation	There are two types of variation – natural (common cause) and artificial (special cause). Examples of natural variation include admission and discharge patterns, waiting list initiatives that impact on the demand on a downstream service, staff holidays. Examples of special cause variation include lost patient notes, a patient changed address, a piece of equipment broke down. Aim to reduce the effects of the natural variation so that all patients receive the same standard of care and the flow through the service is maintained. Try not to focus too much resource on identifying the special cause variation. By using specific rules for SPC (see above) you can identify common and special cause variation and determine whether your process is in control.

Additional resources

There is a wealth of material available on the subjects of DCAQ, SPC and measurement for improvement. Listed below are some suggested websites and documents for further reading.

Demand, Capacity, Activity, Queue (DCAQ)

You can download the computer models used in the masterclass from Richard Steyn's website. There are also several presentations that have been used by his team when delivering their lectures around the world.

http://www.steyn.org.uk/

The Improvement and Support Team (IST) have developed a Continuous Improvement Toolkit which can be accessed from the e-library website below. Select IST and goodpractice.net from the Quick Links on the left-hand side and follow the instructions to log in. From the toolkits page select Toolkit 3.

http://athens.goodpractice.net/nhs/nhs.aspx

IST have also produced other documents relating to the benefits and methods of using DCAQ. Please contact IST for copies at istmailbox@scotland.gsi.gov.uk.

Statistical Process Control (SPC)

The Clinical Indicators Support Team at the NSS Information Services Division (ISD) have put together a website with information and tools to learn more about using SPC.

http://www.indicators.scot.nhs.uk/SPC/SPC.html

Richard's website (see above) also has information about using SPC.

Measurement for Improvement

The Institute for Healthcare Improvement (IHI) website has links to tools, resources and literature so there is plenty of practical information to help both beginners and the more advanced improvement leaders find the best way to measure their process and show where improvements have been made.

http://www.ihi.org/IHI/Topics/Improvement/ImprovementMethods/ImprovementStories/SuccessfulMeasurementForImprovement.htm

The NHS Institute for Innovation and Improvement's No Delays Achiever website has a section dedicated to service improvement.

http://www.nodelaysachiever.nhs.uk/ServiceImprovement/

The Scottish Government

This document is also available on the Scottish Government website:
www.scotland.gov.uk

RR Donnelley B53620 12/07

Further copies are available from
Blackwell's Bookshop
53 South Bridge
Edinburgh
EH1 1YS

Telephone orders and enquiries
0131 622 8283 or 0131 622 8258

Fax orders
0131 557 8149

Email orders
business.edinburgh@blackwell.co.uk

ISBN 978-0-7559-5579-4

9 780755 955794

w w w . s c o t l a n d . g o v . u k