

This issue and archive at <http://www2.wmin.ac.uk/coiec/nosokinetics.htm>  
 Current screen-edited version at <http://www.iol.ie/~rjtechne/millard/index0.htm>  
 Page 1 Pressure and force meets space and time  
 Page 2 -4 Understanding diminishing returns in the reduction of waiting times  
 Professor Michael Pidd, University of Lancaster, UK  
 Page 5-7 Applicability of the OM triangle to health care: Part One  
 Professor Glen Schmidt, University of Utah, Salt Lake City  
 Page 8-9 Selected articles, job opportunities, forthcoming conference



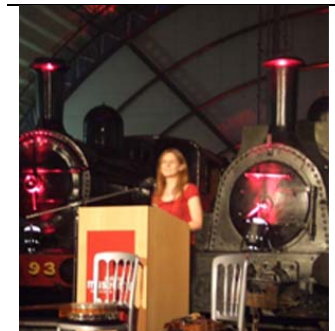
## Pressure and force meets space and time

What will health care look like twenty years from now? Will the high pressure storms that are blowing through health and social care systems blow themselves out? If so, what will replace them? Nigel Hawkes, Health Editor of The Times, writing in the BMJ (23 September 2006; 333: 645-8), gives a different perspective on the reform process. Citing Lenin he states what really matters is “Who has the power over whom? Who is the master and who is the servant.”

Put simply, ‘... what really matters is who does the kicking and who is being kicked.’ Seen in this light, the real purpose of NHS reforms e.g., independent treatment centres, patient choice, payment by results, practice based commissioning, incentives to general practitioners, foundation trusts, targets etc, is to undermine the power of professionals and the unions to make medicine a ‘creature of the state’.

One may well ask, “Given that’s the goal, where do we fit in?” What place do mathematical models have in a world of rhetoric and spin? At the recent Royal Statistical Society conference in Belfast, so competently organised by Adele Marshall, Prof Adrian Raftery a statistical expert in weather forecasting, stated ‘because of the complexity’ it is impossible to forecast weather more than 23 days in advance. Are we better than that?

In this issue we are privileged to have two complementary Professorial contributions which throw light on the complex interactions between supply and demand.



Adele opening conference dinner in railway museum

Using a simple queueing model, Prof Michael Pidd, from the University of Lancaster explains why NHS care Trusts may have extreme difficulty in meeting their 2008 waiting list target of 13 weeks from referral to treatment. Most Trusts have succeeded so far, from 65 weeks to 26 weeks for outpatient appointment, probably by tackling long waits. But, further gains will be harder to achieve because straight lines are meeting curves.

On a similar vein, Prof Glenn Schmidt from the University of Utah uses the OM triangle to explain the interplay between inventory, capacity, and information and the ‘curse of variability’. He also uses the triangle to explain choices in purchasing a MRI scanner. In the next issue he discusses management options in three different clinical scenarios.

Without a coherent long-term plan, can anything be achieved? Today (28<sup>th</sup> Sept 2006) The UK Prime Minister, Tony Blair, in his last year has thirty new projects he wants to introduce before he goes: further modernisation of the health and social care system, reforming education, solving world crises etc. Our ambitions are much simpler; a conference in Ireland in 2008, a revamped newsletter in 2007 and further development of the [nosokinetics.org](http://nosokinetics.org) website.

## Understanding diminishing returns in the reduction of waiting times

Professor Michael Pidd, Department of Management Science, Lancaster University, Lancaster LA1 4YX [m.pidd@lancaster.ac.uk](mailto:m.pidd@lancaster.ac.uk) Link for [The full text of this article](#)

*Simple queuing models clearly demonstrate that the laws of diminishing returns apply to systems in which service capacity is close to the demand for that capacity. Hence, reductions in long waiting times may be fairly easy to achieve but further reductions get harder and harder. It is important that policymakers and managers appreciate this when using waiting targets as policy levers.*

### Introduction: waiting time targets

Pressure on NHS Trusts to change, using targets and star ratings for both outpatient and inpatient care, is a major plank of the government's health service policy. In 2001, Trusts had to ensure that patients waited no longer than 65 weeks (15 months) for their first outpatient appointment. Since then, the targets have gradually become more severe as shown in Table 1.

Waiting time target (weeks)	2001/02	March 2003	March 2004	Dec 2005	Dec 2008*
	65	52	39	26	18/13

\*NB. The December 2008 target of 18 weeks from referral to admission is equivalent, in many cases, to 13 weeks if expressed in terms of the earlier targets. Hereafter, the 2008 target is referred to as the 18/13 week target.

Using a variety of strategies, most English NHS Trusts achieved the December 2005 target of 26 weeks[1]. However it is unclear whether this has been achieved at the expense of aspects of hospital care which are not being measured under the target regime. Will it to be possible for them to achieve the 18/13 week target? Clearly, different Trusts will go about this in different ways but it is worth investigating what such a reduction by 2008 implies in general terms.

This paper uses an extremely simple queuing theory model to point out the processing improvements that are needed if Trusts are to meet their 2008 target. Queuing theory is not being used to give precise numerical estimates, but rather to indicate the order of changes that might be needed to meet these targets. It is intended as a contribution to debate on this subject rather than being a definitive answer.

### The basic model

The simple integrated black-box system shown in Figure 1 is far too simple to be representative except in a very crude way, nevertheless, it can produce some useful insights. Queuing theory enables us to estimate aspects of the performance of a queuing system based on the relationships between the rate at which customers arrive and the rate at which they are served. Queuing theory has been widely used to understand the operational behaviour of detailed aspects of health care provision [2] and for detailed modelling of waiting lists and their impacts[3], [4], [5] but here it is used to reflect on current macro policy.

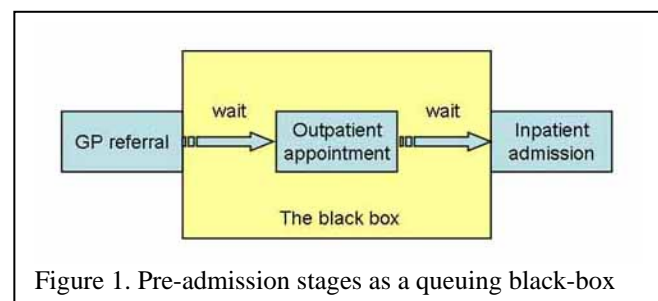


Figure 1. Pre-admission stages as a queuing black-box

The simplest possible queuing system has a deterministic rate of arrival and a deterministic service time: that is, the system literally runs like clockwork. In such systems it is very easy to work out how long a customer will have to wait. Slightly more complicated, but still very simple, is a system in which customers arrive at random and are served at random. Such a

system can be modelled as a so-called M/M/1 queue, which has theoretical properties that are well understood. In such queues the most important property is the ratio of the arrival rate to the service rate, which is known as the traffic intensity. Clearly, if the traffic intensity is greater than one then people are arriving at a faster rate than they can be served, which will result in ever longer queues and waiting times. The lower the traffic intensity, the less time customers are expected to spend in the queue.

Using simple queuing theory for an M/M/1 queue, we can construct Figure 2, which shows the expected time that a customer will wait in the queue given different traffic intensities. The important feature of Figure 2 is the shape of its curve, which shows an exponential decline in the waiting time relative to a linear increase in the processing rate. Hence, when the service rate is close to the arrival rate but starts to fall, there is a very rapid decrease in the expected waiting time relative to a small increase in the service rate. As we move to the right of the curve it gets much flatter, which implies that the relative improvement in waiting times is much less once the traffic intensity is at a lower value.

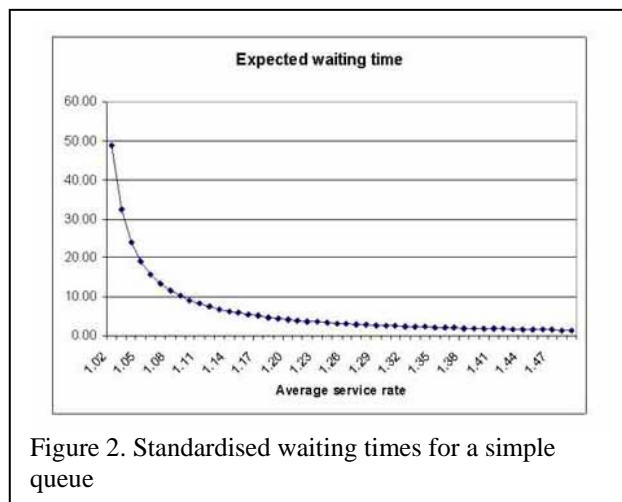


Figure 2. Standardised waiting times for a simple queue

There are only two ways in which the traffic intensity can be decreased: either by reducing the rate at which customers arrive or by increasing the rate at which they are processed. If the arrival rate of customers is not within the control of those operating the queuing system, then their only option is to increase the rate at which customers are served - if they wish to reduce the average waiting time. The simple queuing model makes it clear that there are diminishing returns on subsequent investments required to increase the processing rate; that is, investments in increased capacity.

### Increasingly severe waiting time targets

The waiting time targets in table 1 decrease by 13 weeks on each occasion. This is a linear decrease in waiting times, which is something of a contrast with the shape of the curve shown in figure 2, which has an exponential decline as a result of a linear decrease in the traffic intensity.

Using queuing theory based on this simple queuing model we can also calculate the service rate that would be needed to produce a particular waiting time, given a specified arrival rate. If for simplicity, we standardise the arrival rate as 1 per time unit, Figure 3 shows that reducing the average waiting time from 65 to 25 time units requires a relatively small increase in the service rate. However, further reductions in the average waiting time hit the much steeper part of the curve. So, it is much harder to reduce the average waiting times further, once they have already reached the value of 25 time units.



Figure 3. Relative service rate needed to reduce waiting times

### Discussion

The queuing analysis presented here is far too simple to be translated directly into an analysis of waiting times for inpatient treatment in NHS Trusts. However, it does demonstrate very clearly

the law of diminishing returns which applies to all queuing systems: it takes a lot of effort to squeeze the last drops of water from a towel. It is obviously true that different specialties will adopt different approaches to meeting the waiting time targets; however, the same law of diminishing returns applies once a Trust sets about reducing average waiting times. The simple queuing model shows that relatively small increases in service rates can lead to significant reductions in average waiting times, if a Trust's current performance sits on the flat part of the curve. If the Trust is performing on the steeper part of the curve, then further reductions in average waiting times may be very hard to achieve without further significant investment, or without diverting resources from activities not covered by the targets. It is impossible to use this model to get sensible numerical estimates for any particular Trust since we do not know where the Trust sits on the curve of figure 3. Nevertheless, when setting policy, it is important to understand that the law of diminishing returns will apply. If demand rates are increasing, then resource demands may be even more severe than those implied by the exponential curves shown here.

It is also important to realise that the model presented here uses average waiting times, but the targets refer to maximum waiting times. In statistical terms, the targets relate to the upper tail of the distribution of waiting times, rather than to its mean. Much of the success achieved by Trusts in meeting the targets to date may be because they have removed the longer waits; that is, they have worked on the upper tails of the waiting time distribution. This will reduce the variance of the waiting times but may have a much smaller effect on the mean waiting time. Figure 3 shows the increase in service rates needed to bring about reductions in the mean waiting times. If Trusts have concentrated on the tail of the distribution there still may be some scope for further concentration on the longer waits, however it may now be necessary for them to focus on reducing average waiting time. How easy this will be, depends on where they can be placed on the curve of figure 3.

Finally, waiting lists act as a form of buffer between the demand for services and their provision. In many specialties, demand is seasonal; higher at some periods of the year than others. This allows service managers to provide capacity to meet average demand, knowing that waiting lists will increase at times of high demand but will reduce at times of low demand. Were a service required to have a very short waiting list, it would need to have enough capacity at times of high demand, which means that there will either be spare capacity at times of low demand, or capacity will have to be provided on a seasonal basis. In a cash-limited service, providing spare capacity everywhere seems extremely unwise and it is unclear where extra seasonal capacity could be found. Hence, it is sensible to keep waiting lists just long enough to cover this seasonal variation.

### **Acknowledgements**

I am grateful for comments and suggestions received from Prof Gwyn Bevan & Dr Alec Morton (LSE), Prof Peter C. Smith (University of York) and Dr David Worthington & Murat Gunal (Lancaster University).

### **References**

1. Lewis R, Appleby J. Can the English NHS meet the 18-week waiting list target?. *J R Soc Med.* 2006 Jan;99(1):10-3.
2. Preater J. A bibliography of queues in health and medicine.  
<http://www.chcm.ubc.ca/comm682/QueueingBib.pdf>
3. Worthington DJ. Hospital waiting list management models. *J Operational Res Soc* 1991; 42: 833-843.
4. Worthington D.J. Queueing models for hospital waiting lists. *J Operational Res Soc* 1987; 38: 413-422.
5. Bevan R.G. Taking equity seriously: A dilemma for government from allocating resources to primary care groups. *BMJ* 1998; 316:39-42 (3 January).

## APPLICABILITY OF THE OM TRIANGLE TO HEALTH CARE: Part One

Professor Glen Schmidt, David Eccles School of Business, University of Utah, Salt Lake City  
 glen.schmidt@business.utah.edu

*Editor’s comment: Searching the internet I found the OM (Operations Management) triangle. Here Prof Schmidt explains the interactions between inventory, information and capacity and triangle and describes the curse of variability. [Full text on web.](#)*

**If you are a pilot, you know about the Bermuda triangle. If you study relationships, you have heard of a love triangle. If you enjoy recreational mathematics, you are probably intrigued by Pascal’s triangle. But if you read Nosokinetics News, the triangle you should definitely know about is the OM triangle!**

### OPERATIONS MANAGEMENT

Operations Management (OM) is a field within business and engineering that addresses the flow of materials through production processes. A manufacturing plant takes raw materials and performs certain process steps to transform these raw materials into finished goods. While being processed, the unfinished products are called work-in-process. Thus the firm’s *inventory* is comprised of three components: raw materials, work-in-process and finished products.

Analogously a medical facility has inventory (of patients) just as the factory has inventory (of products) – the only real difference is that in a medical facility you can’t store finished goods to sell to new arriving customers (wouldn’t it be great if you could store a “cured patient” on the shelf and hold it as an immediate cure for a new arriving patient!).

<i>Factories</i>	<i>Medical facilities</i>
Raw materials	Waiting Patients
Work-in-process	Inpatients
Finished goods	Discharged patients

Both the manufacturing plant and the medical facility need to determine the type and number of resources needed in order to transform raw materials into finished goods. For the factory, the resources might consist of presses or machining centers, along with workers. For the medical facility, the resources might be beds, MRI machines, and of course doctors and nurses. The more resources the manufacturer acquires, the more *capacity* it has

Factories and medical facilities are also similar in that they must deal with “the curse of variability.” There are two keys types of *variability*; namely, variability in the arrivals of customers (e.g., the arrival of patients to the hospital), and the variability in processes (e.g., it may take longer to diagnose one patient as compared to another). Both of these types of variability wreak havoc in a production process.

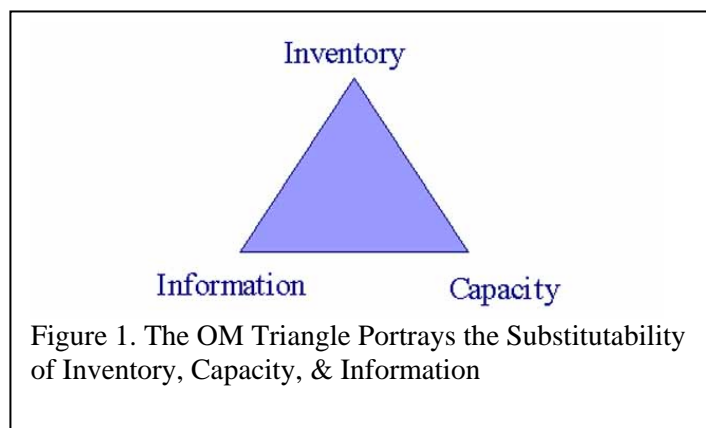


Figure 1. The OM Triangle Portrays the Substitutability of Inventory, Capacity, & Information

How can you keep your doctors and nurses busy, but not overworked, if you don’t know how many patients will arrive that day, or don’t know how long a patient will take to be diagnosed? What each firm would like is *information* as to the extent of this variability and how to reduce it.

Thus a medical facility is similar to a factory in many respects, and the notions of “inventory,” “capacity” and “variability (or information)” can be analogously defined in both settings.

Furthermore, much of the theory that developed from the study of manufacturing can also be applied to services such as health care. Just one example of this is the way some hospitals are implementing the Toyota Production System (Connolly 2005).

Another similarity between manufacturing and services is that they both experience tradeoffs depicted in the “OM triangle” (Lovejoy 1998, Schmidt 2005). See Figure 1. The OM triangle depicts the tradeoff between the three factors described above, namely, inventory, capacity, and information. The theory behind the OM triangle is rigorously grounded in queueing theory – see Schmidt (2005) for a somewhat more formal development.

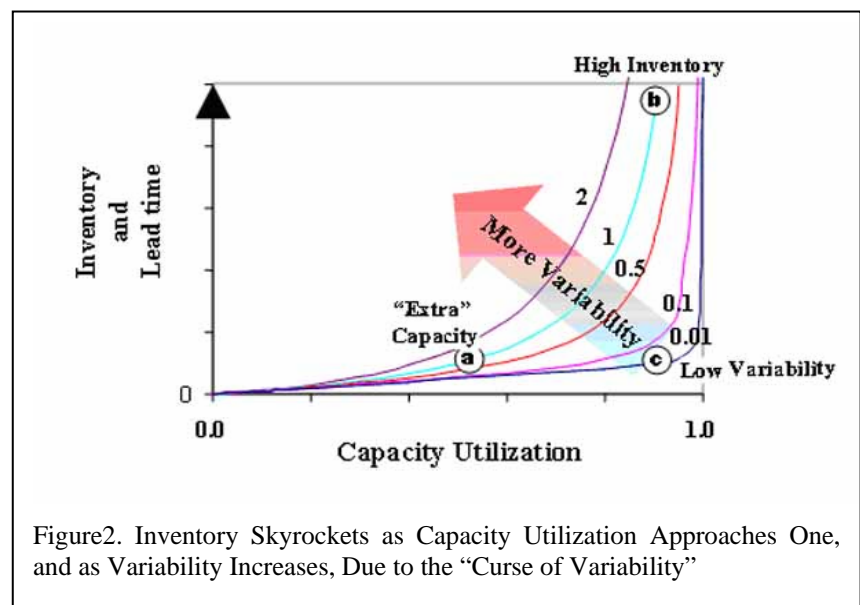
### BUYING AN MRI MACHINE: THREE OPTIONS, WHICH CHOICE?

To illustrate the tradeoff, consider the following example. Say you are in charge of running a small hospital. You are currently looking at buying the hospital its first and only MRI machine. Assume you know that on average a patient will arrive every hour or so, but not necessarily with perfect regularity. A less expensive MRI machine can examine a patient in 50 minutes on average, while a more expensive one only takes 30 minutes (but again, there is some variability in the processing times, as not all patients need the same scans). You can only buy one – which one should it be?

Here you are making a decision on how much *capacity* to acquire. The expensive machine has the capacity to examine two patients per hour, while the less expensive one can examine one every 50 minutes which is equivalent to a lower processing rate of only 1.2 per hour.

Recall that patients arrive at 1 per hour, on average. Say you buy the higher-capacity machine; the one that can process 2 patients per hour. Even though you will be using the machine only half the time (i.e., its capacity utilization will be 30 min. / 60 min. = 0.5 since it can treat a patient in 30 min. and a new patient arrives every 60 min. on average), a patient will sometimes have to wait – sometimes the next patient will arrive before the current patient is done.

A second option is to buy the lower-capacity machine. You will still have more capacity than the minimum needed (its capacity utilization will be 50 min. / 60 min. = 0.83). However, in this case the patients’ average waiting time will increase significantly over the wait time experienced with the faster machine. The average number of patients waiting, i.e., the ‘raw material inventory’, will also increase significantly. This dramatic increase is due to “the curse of variability” – if the capacity or rate at which patients can be treated is not much greater than the rate at which they arrive, variability in the system drives the inventory “sky-high” (and the more variability there is, the worse this relationship gets). See Figure 2.



The above discussion implies there is a relationship between capacity and inventory. Thus you can think of your decision as one of how much *inventory* to hold, rather than how much capacity to acquire. Just as buying the machine with more capacity costs you more money, holding more inventory also costs you more. The cost of holding more raw materials (that is, of having customers wait) is not an out-of-pocket cost, but instead it is the less tangible cost of having

disgruntled customers. You need to think about the degree of dissatisfaction that customers will experience from having to wait, and whether this is tolerable.

A third alternative is to buy the less expensive machine while also spending some money on *information*. For example, you might invest in an information system that tells you exactly what medical tests every patient in the hospital needs at every point in time. With this information, you might be able to develop a scheduling and tracking system that minimizes each customer's wait, and thereby effectively reduce the variability in arrivals to the MRI machine. In turn, reducing variability in arrivals would reduce the average number of people waiting.

### **Three alternatives**

In the above discussion we described three alternatives on how to “spend your money.” To reiterate:

1. Option one is to pay the high cost of the fast machine, which minimizes waiting lines and waiting costs.
2. Option two is to reduce machine cost by buying the slow machine, but this will cost you in terms of customer dissatisfaction.
3. Option three is to buy the slow machine but simultaneously buy the information system, thereby again minimizing waiting times and avoiding the cost of dissatisfaction.

As this discussion suggests, each option has its costs – it is your job to determine which alternative has the lowest total cost. The option with the lowest total cost would be your best choice (assuming these are the only relevant considerations).

Said another way, capacity and inventory and information are substitutes. If you acquire more capacity you can get by with less information or drive inventory (and customer dissatisfaction) down. If you hold more inventory, you can get by with less capacity or information. If you acquire more information you can get by with less capacity or drive inventory down.

(to be continued in December issue or on [Full text on web.](#))

*NB The copyright owner for this paper is Prof Glenn Schmidt.*

## **Measuring Productivity in Health Service Delivery an Australian Government Initiative.**

Mark Mackay drew to our notice, that the Australian Commonwealth Treasury Department's has asked the Australian Productivity Commission to undertake a feasibility study into the prospects of estimating productivity growth in the area of health service delivery and, if practicable, to prepare experimental estimates of health sector productivity. For those who have worked on modelling issues in Australia, this research may be worth watching.

Contributions are being arranged via State Treasury Departments although non-Government researchers may wish to approach the Commission directly. The expected release date of the project is Nov 2006. Contact: Paul Gretton (02) 6240 3252, Assistant Commissioner, Trade and Economic Studies. <http://www.pc.gov.au/researchproject/2005/051102.html>

## Readmission

Billings J, Dixon J, Mijanovich T, Wennberg D. **Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients.** *Bmj* 2006;333(7563):327.

Uses multivariate statistical analysis to model the chance of readmission, based on a 10% sample of English HES data (1999-2004) specified conditions known to be associated with predictable readmission (e.g. congestive heart failure, diabetes, COPD, sickle cell disease). Algorithm score 0-100 based on 21 most powerful – statistically significant – variables. Tested model on 10% sample. Key variables were age, sex, ethnicity and number of previous admissions. At a score of 50, 54.3% of readmissions identified correctly, 34.7% flagged incorrectly. At 70 and 80, 22.6% and 15.7% flagged incorrectly. Paper also describes regional differences in readmissions.

## Failure time analysis

Vollmer RT. **Analysis of turnaround times in pathology: an approach using failure time analysis.** *American Journal of Clinical Pathology* 2006;126(2):215-20.

Uses failure time analysis methods to study turnaround times in pathology. The model views a laboratory specimen like a living patient. When the specimen enters the laboratory, the time is analogous to the time of diagnosis for a patient. When the specimen's analysis is completed, the event is analogous to a patient who has died. Shows that the Kaplan-Meier plotting method, the log-rank test, and the Cox model can all be applied to turnaround times and provide useful results.

**Pragmatism:** OED Philos. The doctrine that the whole meaning of a conception expresses itself in practical consequences 1898.

Ormerod R. **History and ideas of pragmatism** *Journal of the Operational Research Society* 2006;57:892-909

Lists 12 reasons for thinking that pragmatism could serve OR practitioners well

1. Pragmatism is what we do, how practitioners behave
2. Pragmatism supports an empirical (in other words scientific) approach
3. Pragmatism emphasizes the uncertainty and changing nature of findings
4. Pragmatism recognizes the individual psychological nature of meaning
5. Pragmatism holds that inquiry is social, as is knowledge
6. Pragmatism supports a theory of learning based on experience, experimentation and action
7. Pragmatism addresses morality, social interests and politics
8. Pragmatism places theory in the service of knowledge
9. Many OR approaches can find support in the philosophy
10. Pragmatism's stance on many things seems surprisingly modern
11. Pragmatism's biological approach should stand it in good stead to adapt to new science
12. Pragmatism is flexible enough to accommodate other philosophical positions.

## Baffled by statistics: [try this](#)

Joseph M Civetta: **Statistics, The Literature, Hospital Data and Patient Profiles: A Survival Guide: The Internet Journal of Anaesthesiology. 1999; Volume 3, Number 4.**

Thierry Chausalet drew our attention to this paper. Here you will find all you need to know. Given a “statistically significant” result, does it have clinical or biological significance? How do you tell the good from the bad? As poor study design, chance, confounding variables and bias in observation and conduct can influence the results. Whatever your role whether manager, clinician or researcher, this paper (freely available on the web) is a classic.

**Academic Positions in Statistics and Operational Research  
at Queen's University, Belfast, Northern Ireland**

School of Mathematics and Physics

*Centre for Statistical Science and Operational Research (CenSSOR)*

**Readership in Statistics / Operational Research (Full-time)** (£41,133 - £52,106) per annum)

**Lecturerships (2) in Statistics / Operational Research (Full-time)** (£28,849 -42,367)

CenSSOR is a vibrant research centre, whose members have nurtured a strong research environment and team spirit. Current research interests: Survival Analysis, Bayesian Networks, Markov Modelling and Stochastic Models. One focus of research application is the development of healthcare research and medical SOR

Successful candidates will be expected to develop their own research plans, further enhancing the research in the Centre, while demonstrating a commitment to teaching excellence at all levels in SOR.

**Closing date: Friday 29<sup>th</sup> October 2006.**

**Further information and to download application pack -[www.qub.ac.uk/jobs](http://www.qub.ac.uk/jobs)**

**University of Westminster, London  
Research Associate: 1 Year, Fixed Term**

To undertake a 12-month study describing the care and administrative processes involved in the hospital care for older patients with hip fracture. Funded by SPARC (Strategic Promotion of Ageing Research Capacity) commencing November 2006.

Salary Dept Location £23,742 p.a. (Inc. LWA) HSCS Research Harrow Campus

**Closing date: 9th October 2006; Interviews likely to be held on: 20th October**

Informal enquiries about the position to Dr Christos Vasilakis by e-mailing  
[C.M.Vasilakis@westminster.ac.uk](mailto:C.M.Vasilakis@westminster.ac.uk) or by telephoning 020 7911 5000 ext. 4029.

**Forthcoming Conferences**

**5th IMA QUANTITATIVE MODELLING IN THE MANAGEMENT OF HEALTH CARE**

Goodenough College, Central London on 2nd - 4th April 2007

[Conference website](#) / or the IMA website <http://www.ima.org.uk/>

Abstracts of 300-500 words to Lucy Nye at [Lucy.Nye@ima.org.uk](mailto:Lucy.Nye@ima.org.uk) by 1 DECEMBER 2006.

Authors of accepted abstracts will be notified by 1 January 2007.

Selected papers presented at the conference (whether orally or as a poster) will be published in the Springer journal Health Care Management Science.

Dr. T.J. Chausset, Reader, Department of Information Systems, University of Westminster, 115 New Cavendish Street, London W1W 6UW. Email: [chausst@wmin.ac.uk](mailto:chausst@wmin.ac.uk)

Nosokinetics News is mailed to supporters and collaborators interested in developing a scientifically valid approach to measuring and modeling health and social care systems. To be added to / removed from the mailing list email Prof Peter Millard [nosokinetics@tiscali.co.uk](mailto:nosokinetics@tiscali.co.uk)  
For earlier editions <http://www2.wmin.ac.uk/coiec/nosokinetics.htm>  
The on line version is at <http://www.iol.ie/~rjtechne/millard/index0.htm>.  
<http://www.nosokinetics.org/>