Proceedings of the 2004 Crystal Ball User Conference

FINANCIAL RISK IN HEALTHCARE PROVISION AND CONTRACTS

Dr Rod Jones (ACMA, CGMA)
Healthcare Analysis & Forecasting
Camberley, Surrey
Tel: +1276 21061, Mobile +7890 640399
hcaf_rod@yahoo.co.uk

For further articles on financial risk in healthcare go to: www.hcaf.biz

ABSTRACT

The provision of (i.e. resource allocation issues) and payment for (i.e. financial risk issues) healthcare relies on the volume and mix of conditions treated. This paper will explore the nature of statistical distributions behind healthcare demand and will explore the impact of environment (e.g. weather, viruses, etc) on the variation in demand. The financial risk arises from the variation in demand times the price for service/treatment. Simulation can be used to explore the risk associated with various forms of healthcare contracts. Examples will be given using UK HRG (similar to US DRG) payment.

1 SUMMARY OF KEY POINTS

The Concept of Randomness and Variation

1.1. Contrary to our subconscious assumptions demand is not constant. Demand is variable which makes averages an unhelpful planning tool.

1.2. There are 2 sources of variation. These are:

a) That which is caused by the fact that healthcare demand operates within a complex system of short and long term cycles which means that the average is changing over time (special cause variation).

b) However, even if we knew the true average over time our actual ability to measure it – and deliver services to that level - is obscured by the fact that there is statistical variation around that average (common cause variation). This statistical based variation is described by Poisson Statistics. Hence the standard deviation (a measure of variation) is equal to the square root of the expected average

1.3. In practice the expected average is no longer accurately known because it is obscured by this statistical randomness.

What Can Be Done?

1.4. The implications for planners and operational managers is that we have to start to apply ranges with upper and lower limits rather than pretending that we know the true and precise value of demand.

1.5. Healthcare demand is actually quite small when examined on a daily or weekly basis and when split down to clinician level within an individual specialty or service. Demand on a daily, weekly or monthly basis is therefore so uncertain that the average loses its meaning for resource allocation and staff on the ground. This means that the variation is very high in percentage terms – and it is this phenomenon which creates a sense of lack of control.

1.6. The way to start to solve this is by the introduction of upper and lower control limits coupled with industry type control charts to complement the management process. In addition there is a need to develop approaches which enables the more flexible use of resources across teams, wards, etc.

1.7. Attempts to service variable demand using services based around an average will lead to the formation of queues as witnessed in A&E departments, and UK outpatient and inpatient waiting lists. This is only made worse when the real capacity is lower than the presenting demand. This means that optimum efficiency is actually achieved with slight over capacity (but with the assumption that staffing levels will be flexed to minimise revenue costs)

1.8. Queuing theory (which is based on Poisson Statistics) and simulation can be used to help understand the resource allocation issues.

Financial Risk in Health Care Contracts

1.9. To the purchaser the financial risk can be contained by setting a single value contract to cover provision of health care to the population, i.e. all the risk is moved to the provider of services or by contracting at cost per case and attempting to regulate demand, i.e. risk is offset by reducing demand

1.10. To the health care provider the financial risk is a complex mixture of income (fixed or variable) and costs (with associated random variation in both fixed and variable costs!)

> Unless you understand the principle of variation in demand you will never understand healthcare resource allocation or financial risk.

2 THE NATURE OF HEALTHCARE DEMAND

The planning of healthcare services be they buildings, equipment or human resources or the contracts/financial calculations supporting these services will usually start with some estimate of future demand. Such demand estimates are usually made at the expected average where it is implicitly assumed that variation around the average is relatively small and that if it does exist such variation will be symmetrically distributed (as in a normal distribution) around the average. Hence as a provider of healthcare services our assumed income per type of treatment is simply average demand times price and to derive our profit we subtract fixed costs and demand times variable costs, i.e. Profit = Demand x (price - variable costs) – fixed costs.

Is this a valid assumption or do we need to delve a little further? Figure One gives the results of many hundreds of studies into the level of variation associated with many types of healthcare demand. In this figure the X-axis gives the average volume of demand in the period of interest (day, month, year, etc) while the Y-axis gives the observed or apparent standard deviation associated with this average. The apparent standard deviation makes no assumption regarding the type of statistical distribution but simply applies the textbook formula for calculating the standard deviation of a normal distribution.

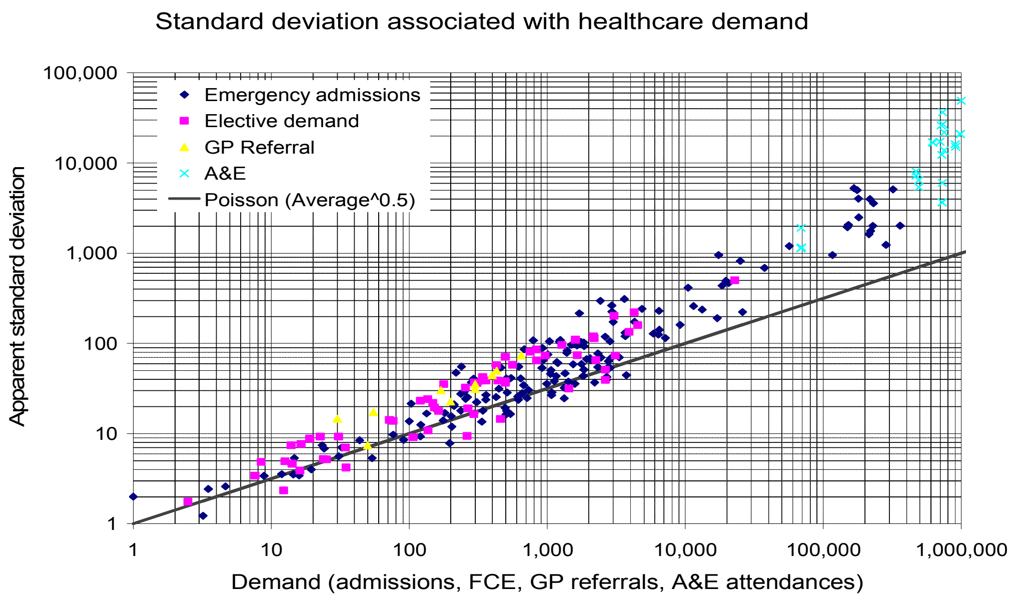

Figure One: Natural level of variation around the average for different types of healthcare demand.

In the discipline of statistical process control the standard deviation is said to be 'the voice of the process', i.e. every process has an associated statistical signature. However in most cases the standard deviation is not assumed to be directly linked to volume in the way that healthcare demand appears to be. The reasons behind this linkage can be explained by the combined effects of 'common cause variation' and 'special cause' variation.

2.1 Common cause variation – statistical variation around the average

For any given average there will always be variation around this average due to statistical-based randomness. For most healthcare demand this type of variation is described by Poisson randomness.

Poisson statistics describes arrival/demand for service events such as telephone calls per hour at a switchboard, customers per hour, GP referrals per week, emergency admissions per day, etc. The outcome can only be an integer value (i.e. we had 10 GP referrals last week) although the expected average can be a decimal value (i.e. our average is 8.4 GP referrals per week).

One highly interesting feature of Poisson statistics is that the standard deviation around the average is always equal to the square root of the average – hence the straight line in Figure one. Unlike the Normal Distribution where the spread of events around the average is symmetric that of a Poisson distribution is skewed. Hence there is a tendency for more events with a value less than the average but with a tail of infrequent events at much higher than the average. This tail causes havoc to healthcare services. In percentage terms the standard deviation associated with healthcare demand will therefore be intrinsically high which is then the fundamental basis for financial risk and the required size of service to meet demand..

2.2 Special cause variation

However it is clearly seen from Figure one that most healthcare demand has a standard deviation which is larger than that predicted by simple Poisson variation. This is partly a product of the fact that the average changes over time and that the measured or apparent standard deviation is the result of both special and common cause variation.

2.2.1 The average is changing over time

- Circadian Cycles – all biological systems show circadian (i.e. 24 hour cycles), hence, for particular conditions the true incidence rate (admissions per hour) will vary with the time of day.
- Daily Working Patterns - this is further complicated by GP working hours and the availability of any other supporting services. The overall effect is a distinct daily cycle in emergency admissions (greatest during working hours) competing with a working hours pattern for elective admissions.
- Weekly Working Patterns – GP's, Social services, Home Care, etc almost all 'work' a five day week less any public holidays and hence GP referrals (including emergency admissions) likewise show distinct working day patterns. For some specialties emergency admissions peak on a Friday (prior to the weekend) and a Monday (after the weekend).
- Seasonal Cycles – these are the annual cycles which depend on the type of weather, viruses and other infections prevalent at different times of the year. Also included are the impact of school holidays and consequent flow of large numbers of people to holiday locations.
- Longer Term Cycles – the incidence of particular conditions also appears to follow longer term cycles. These longer-term trends are poorly understood. An example is given in Figure two for a large Trauma & Orthopaedic department in Berkshire (England) where there is a long-term average of 7 emergency arrivals per day. The chart is interpreted by realizing that each point represents the average over the previous 6 months. Hence the peak of 8.1 arrivals per day in September 1995 is an average of arrivals per day over the period April to September, i.e. roughly the spring and summer months. However, in particular years, e.g. 1998 the arrivals during this 6 month period only averaged 6.8 per day. Autumn/winter arrivals are just as variable and can range from an average of 6 to 7.3 per day. Note that the highest number of emergency admissions for a single day was 29 on 30th December, 1995 when melting snow turned to ice, i.e. over 4 times the annual average!

  Figure two also emphasizes the importance of taking the longer term view since if we were to base future forecasts on data from 1998 onward we may be tempted to say there was a trend upward. It is disappointing to note that most NHS organizations rarely have data going back longer that four to five years, i.e. the old style bed planning methodology and indeed the process of contracting with purchasers gave little emphasis to looking at historical trends and hence the data was not valued enough to consider keeping! Indeed the variation caused by common and special cause variation contradicted the accepted view of how healthcare demand should behave and the validity of the data was questioned rather than questioning the heuristic viewpoint.

Figure two also shows the behavior arising from simple Poisson randomness in daily arrivals. As can be seen this can lead to an apparent range in a six-month average of arrivals from 6.5 to 7.5 per day. The apparent co-incidence of some of the peaks is purely an artifact of randomness. It is a curious fact that the outcome of random events usually leads to clusters, i.e. deviation from the average is the norm rather than the exception.

- Population Demography – growth within different age bands will lead to subtle shifts in total healthcare demand.
- Socological & Technological trends – these influence GP referral thresholds and the range of interventions available. For some services such as the breast clinic the rate of referral will be influenced by media coverage and even events within the latest TV 'soap'.
- Step Change – a combination of any of the above factors can lead to a step change in demand. For example a new drug or procedure can create a whole new field of demand. Step changes are exceedingly common in healthcare demand and are usually difficult to predict in terms of arrival and magnitude.

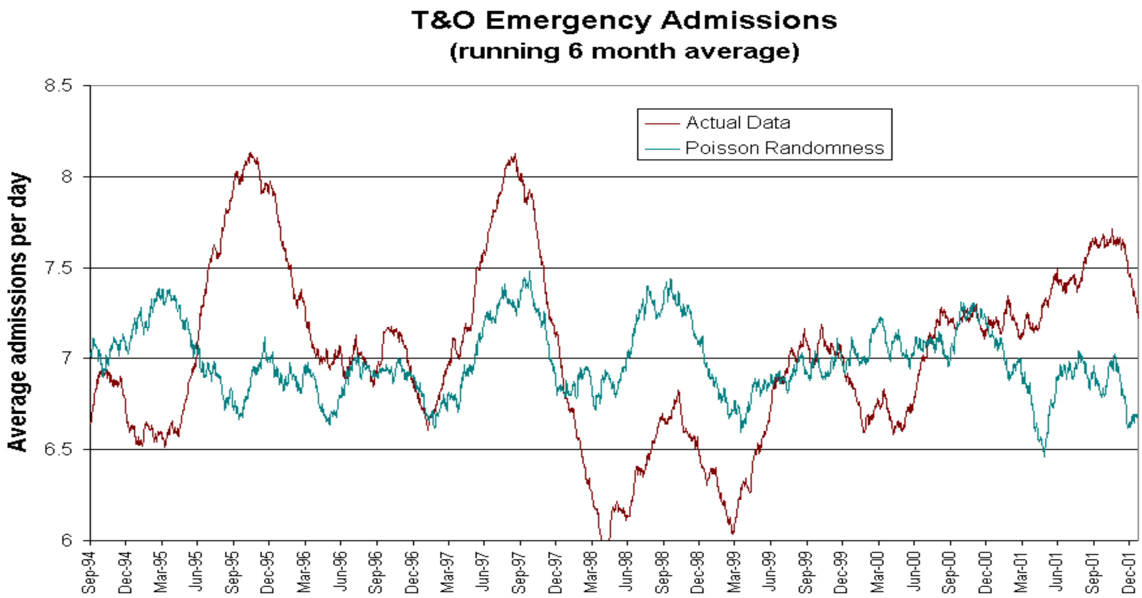

Figure two: Longer term cycles seen in emergency Orthopaedic admission to a large hospital

Given that there are at least 8 broad mechanisms for change in the 'average' you will now understand the need to articulate the exact specification relating to the particular hourly, daily, weekly, monthly or annual average to which you are referring.

You will also immediately appreciate the need for long-term data collection in order to determine the relative effect of the various cycles and trends.

2.2.2 Special and common cause variation are superimposed

The measurement of variation is obviously dependant on the availability of sufficient data which makes it difficult to separate special cause and common cause variation. For example, emergency admission is typically influenced by the patterns of weather and prevailing viral and other infections – as seen in Figure two. The resulting distribution for the annual totals will therefore be a combination of Poisson variation around the constantly moving average plus another distribution(s) describing the movement in the average.

This interaction can be minimized by using 12 month totals, i.e. the underlying seasonal pattern is repeated every 12 months, however, even within the seasonal pattern there will still be considerable special cause events, i.e. an unusually hot summer, an outbreak of influenza A or B in winter, etc. An example is given in Figure three for emergency medical admissions to a large hospital. Experience shows that the resulting distributions are most commonly approximated by a Log-normal, Logistic, Gamma or Extreme value distribution (as per the Batch Fit function within Crystal Ball). All of these distributions are skewed with a long tail of values higher than average or higher than the mode, i.e. the 'common cause' skew implied by Poisson randomness is further extended by the 'special cause' effectors within the environment.

Hence in Fig 3 the maximum value is 2,121 higher than the mode (most common value) while the minimum value is only 1,590 lower than the mode.

The implications of such skewed distributions to healthcare resource and financial planning should be obvious as should the need for risk evaluation tools such as Crystal Ball. The simple accounting assumption of average demand times price = expected income is no longer valid and does not explain the exposure to risk implied by the skewed distributions characteristic of healthcare demand.

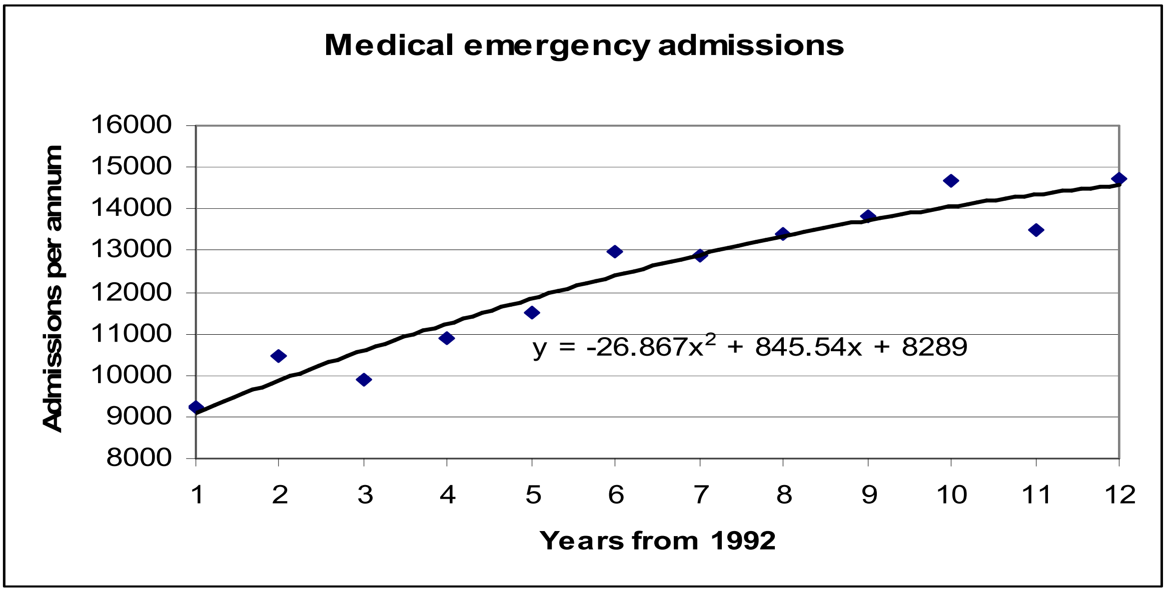

Figure three: Variation in medical emergency admissions over a twelve year period – note the non-linear time trend.

3 THE CONSEQUENCES OF VARIATION

This high intrinsic variation has many consequences to operational and financial planning. These can be summarized as follows:

- The expected future average is no longer known with certainty, i.e. the standard error of the mean is the observed standard deviation divided by the square root of the number of data points. In layman's terminology it is difficult to estimate the average for next year (as per Fig 3) if you only have 2 or 3 years data. Fig 2 and 3 shows false trends, etc.
- It is the year to year variation that drives operational performance rather than the growth – as per Fig 2 and 3.
- At an operational level it is the daily variation that creates the pressures, i.e. 15,000 medical emergency admissions per annum (as per Fig 3) is only 41 per day with 80% of full range between 29 to 53 admissions per day – at this level the average ceases to have meaning – most hospitals are in fact smaller than this example.
- To deliver the average implies a surplus of resource to meet the peaks, i.e. average + 3 x standard deviation gives full range demand – the smaller you are the harder it gets!.
- Traditional financial approaches no longer give correct answers (see Fig 4) since the calculated cost for low and even 'high' volume procedures appears to be driven by variation rather than the real cost. It is very difficult to

allocate non-specific overhead costs to one-off activities (including the artifacts of coding errors). Errors of mis-specification then effect all other HRG/DRGs.

- Simulation tools are therefore vital to help understand the nature of demand and its implications to operational and financial planning.

4 THE NATURE OF VARIATION IN INCOME

The expected income is a function of the type of contract. Hence at one extreme there is the lump sum contract to provide heath care services for a defined population, i.e. all the risk is passed from purchaser (or insurer) to then provider. In this instance the provider needs to understand the nature of fixed and variable costs related with the variable demand. Will the offered income meet the full range of expected costs and leave an adequate operating profit?

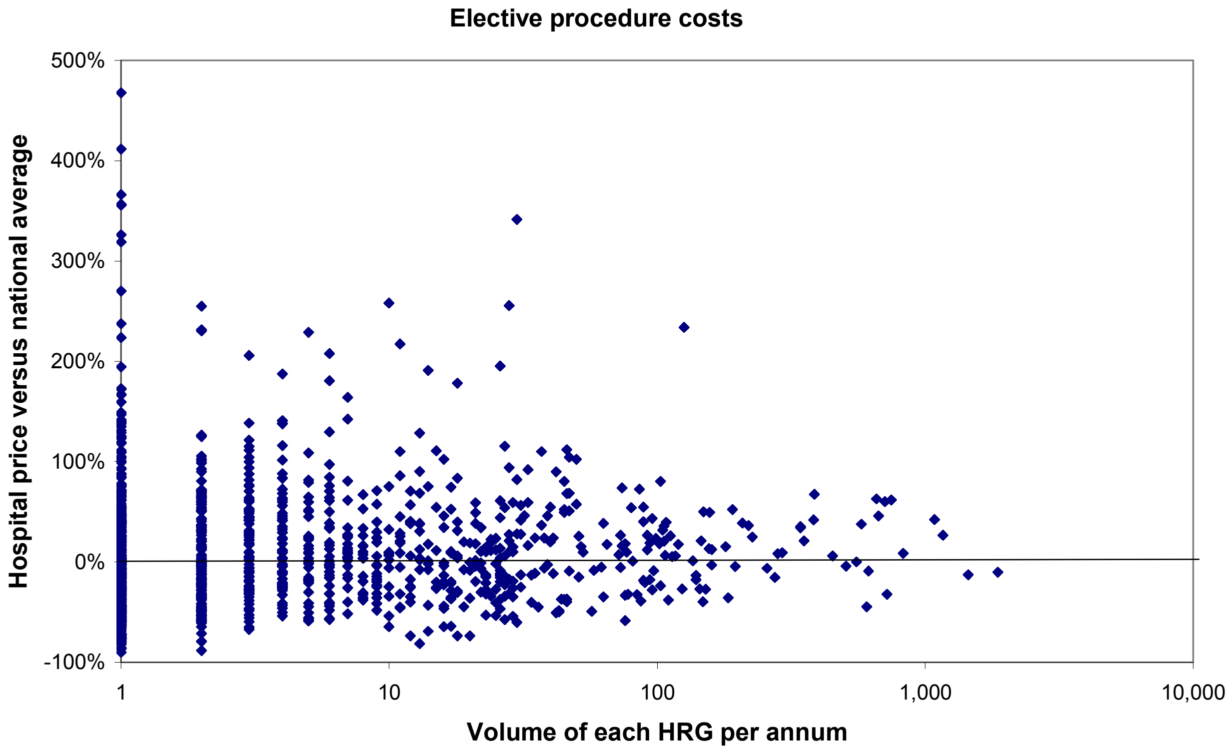

Figure Four: Calculated cost for each individual HRG at a single hospital compared to the national average.

At the other extreme the provider is offered a fixed fee per type of service. The fee is defined by some sort of Health Resource Group (HRG) or Diagnostic Related Group (DRG). For example, a primary knee replacement (HRG H04) may be worth £4,560, etc. In this instance the income is variable and is the sum of case mix time price. For an acute hospital offering a wide rang of outpatient and inpatient (emergency and elective) treatment the resulting income variation can be quite extreme and can only be anticipated by the use of simulation. Figure five gives the expected variation in elective income for a large hospital. In this simulation (using Crystal Ball) at total of 31,140 elective-only FCE spread over 1,234 HRG types were condensed down to 20 price bands each containing around 1,560 FCE per price band. Weighted average price was then calculated for each band. The resulting distribution of income and volume assumed Poisson variation in the volume for each price band. The real world variation will be higher than this, however, the simulation in this instance was designed to be a decision making aide rather than a precise replication of the real world.

5 THE NATURE OF VARIATION IN COSTS

The cost of providing healthcare in its simplest form is derived from the fixed and variable costs. As with all real world processes there is s statistical signature associated with all types of costs be they fixed or variable.

5.1 Variation in fixed costs

Most accounting systems are unable to deal with the unavoidable variation in fixed costs. For example, an insurance premium will have a clause which states that the insured will pay the first $100 or $1,000 of each claim. Hence the fixed cost has a statistical signature which is related to the frequency of claims which in turn the effects next years insurance premium. Salaried staff can likewise be considered a fixed cost, however, there is an associated statistical signature relating to short and long term illness (the cost of temporary staff to cover for absent salaried staff) and the recruitment costs associated with staff turnover. Some staff are likewise more productive than others hence the output after recruitment my be higher or lower, etc.

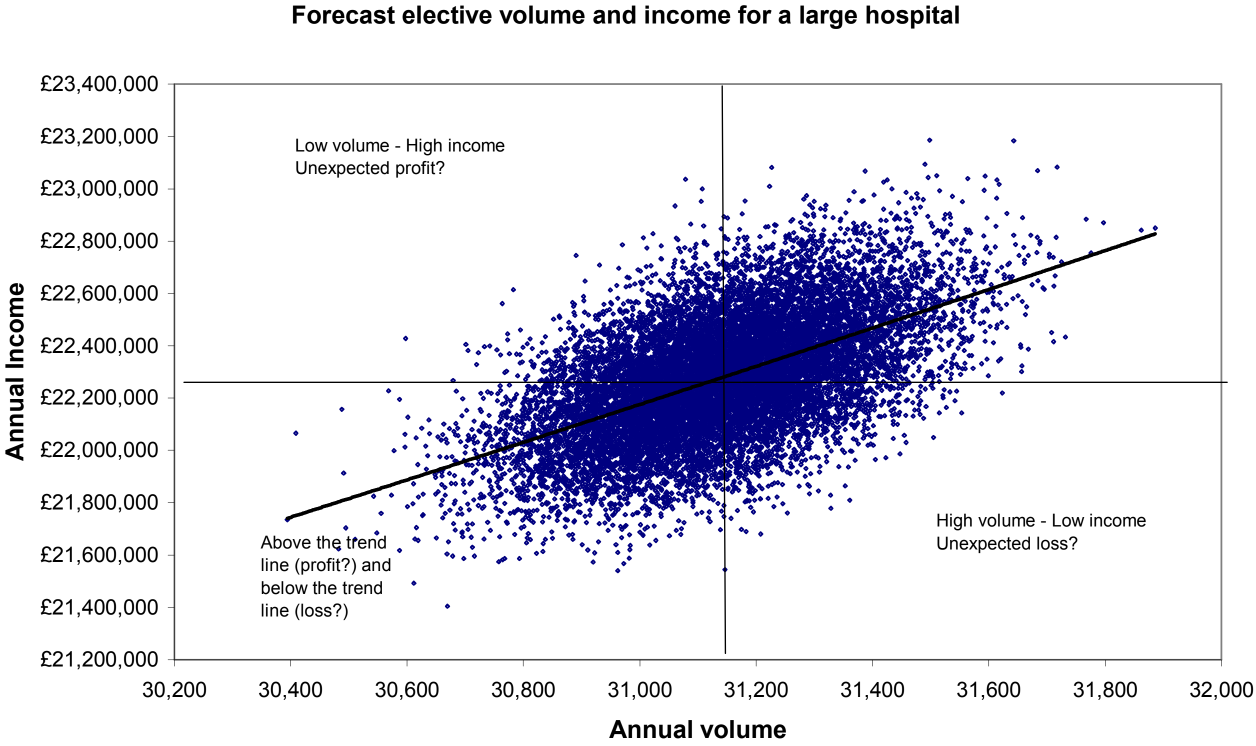

Figure Five: Forecast elective-only annual volume (consultant episodes) and associated HRG-based income for a very large acute hospital. Note that the data is skewed away from the trend line which is driven by the dense cluster close to the average rather than the wider flung values.

The statistical distributions describing the variation in fixed costs also tend to be skewed such as Poisson or related distribution types. The statistical variation in fixed costs is rarely accounted for as part of the contract evaluation or budget setting process mainly because it is impossible to grapple with the implied uncertainty in the absence of a simulation tool. The resulting output is an estimate of the budget or fixed cost contingency allowance. Unless such a contingency allowance is provided the only outcome is an erosion of profits.

5.2 Variation in variable (or volume related) costs

The variable costs are usually those which are patient and procedure related. Payment using HRG/DRG as a basis has the implied assumption that each HRG/DRG is iso-resource, i.e. costs roughly the same for all patients. Figure Six gives one example of what iso-resource means in terms of the variable costs related to length of stay.

Most HRGs are considerably less iso-resource than F82 and hence while the income is fixed at around £1,200 the length of stay related cost can be somewhat variable depending on the volume – high volume hospitals experience less variation than low volume sites due to the sampling effect.

One factor driving the LOS variation is the age of the patient since average LOS increases with age – for obvious biological reasons. Each HRG/DRG has the implied age profile of the national average while the local age profile may be quite different and subject to random variation in arrivals for each age band. It is this age related effect that explains why single year age bands MUST be used to compare relative death rates between hospitals, i.e. local variation from the implied age profile can be considerable and leads to incorrect conclusions. The use of simulation software can be a good defense against allegations of 'high' relative rates of decease following a procedure.

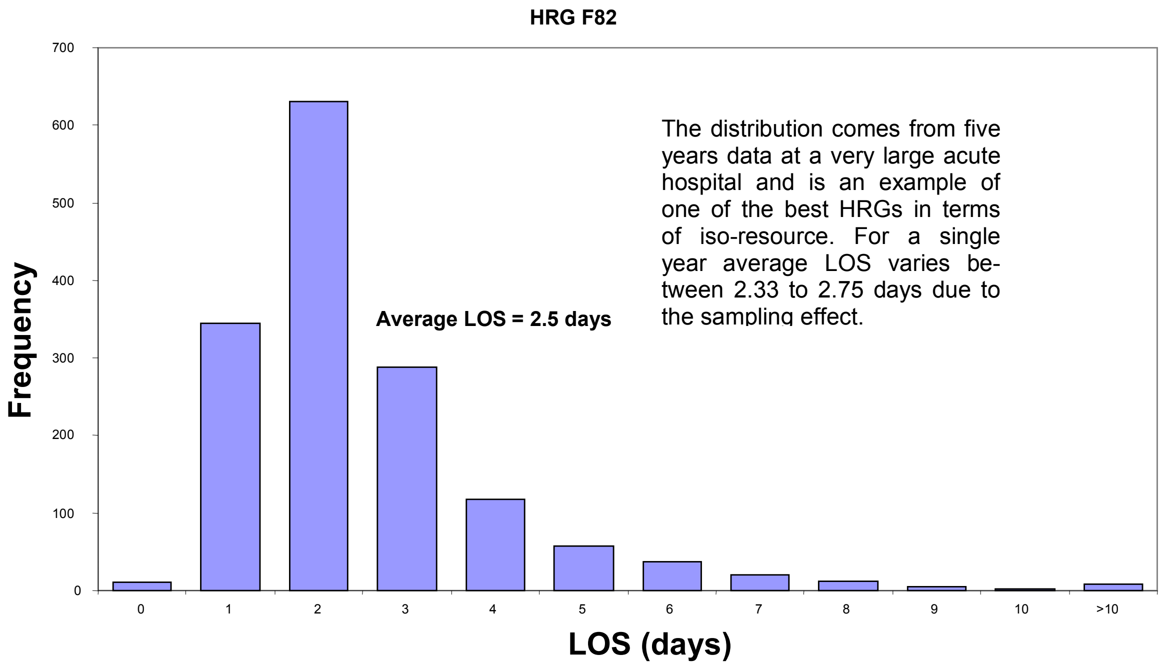

Figure Six: Length of stay variation implied by the use of a HRG/DRG. Random sampling from this distribution determines the total cost for each year. F82 = Appendectomy, patient aged 69 or under and with no complications.

Further sources of variation in the variable costs of each procedure can be traced to three causes:

- The patient – each patient is an individual with associated co-morbidities, hence, consumption of resources such as anesthetic for the operation, post operative pain relief, complications, etc will follow a distribution similar in shape to Fig 6.
- The team – for surgical interventions each surgeon has a different level of skill for each procedure hence theatre time will vary. Post operative recovery will likewise vary depending on the skill of the team supporting the particular pathway of care.
- The hospital – research shows that doctors assume the length of stay of the hospital where they work, i.e. it is the whole system that imposes different rate limiting steps, i.e. the speed of diagnostic support services, etc.

- Accurate clinical coding to enable the range of care to be translated into the most appropriate HRG/DRG. Both Fig 4 and Fig 6 contain hints to the fact that coding errors are in themselves a further source of variation. In this respect the average US hospital is far more advanced than its counterpart in the UK (where a clinical coding administrator starts their career at £12,000 p.a.)
- Errors in the specification of each HRG/DRG. For example, the same procedure for a baby costs far more than for an adult, sub-division of HRGs is probably required for patients with hemophilia or other presenting complicating factors. The national average makes many implied assumptions which are translated into local gains or losses.

6 CONCLUSIONS

Simulation is the only tool available which enables the correct shape of distribution to be appended to the multiple contributors to income and cost. The aim of the simulation is NOT to perfectly replicate the real world but to incorporate sufficient of the real world to enable management to make more informed risk-based decisions.

Hopefully this paper will stimulate hospital finance departments to become more value added and to use appropriate simulation tools to enable their organizations to remain profitable.

BIOGRAPHY

Rod has a Ph.D. in Chemical Engineering and is a chartered accountant (Chartered Institute of Management Accountants). In the early 90's he was among the first group of Crystal Ball users in the UK and has been using the software since that time to solve a wide range of healthcare resource allocation problems.

His career includes 7 years in academia and 10 years in industry covering the biotechnology and food industries and as General Manager of an international laboratory proficiency testing organization. In this latter role he authored a handbook on statistical quality control for microbiological water testing. He has completed the Hewlett Packard intensive course in Total Quality Management. He has over 15 years experience in healthcare both within the NHS and as an independent consultant.

Many thousands of hours of research into healthcare resource allocation has led to the development of new methods for sizing hospitals and bed pools, forecasting healthcare demand, understanding resource allocation problems and predicting the risk associated with HRG/DRG based income/cost. These new techniques are informed by the principles of randomness and variation including the use of queuing theory, computer simulation, and forecasting tools which help to underpin healthcare contracting. The natural level of variation associated with healthcare demand is used to specify contract tolerances, bed pool sizes and to calculate financial risk.

Rod is a regular speaker at conference events and was an invited speaker at the 2003 Health Forecasting Conference hosted by the UK Metrological Office. During 2001/02 he was involved in the UK Department of Health 'Information for Action' good practice guide for the use of hospital operational intelligence to match capacity with demand. He currently provides analysis and advice to Strategic Health Authorities, Primary Care Organisations and NHS Trusts.