A Model for Predicting Healthcare Demand
&
Excess Levels of Utilisation

Dr Rod Jones
Statistical Advisor
Healthcare Analysis & Forecasting
hcaf_rod@yahoo.co.uk or 07890 640399

Key Points

- Appears to be the first model capable of explaining the high levels of unexplained variation observed in previous modelling attempts including the capitation formula

- Model uses the usual predictors of age, deprivation (IMD) & ethnicity

- Main factors which act to explain the unexplained variation are:

    - Students have fewer elective interventions than non-students
    - Population in an acute site catchment area shows unique behaviour
        - This is a reflection of the system behaviours of GPs, Consultants, Ambulance and other supporting services.
        - Emergency admissions show a unique distance dependency for each acute site
        - Each site 'counts' and codes 'admission' events differently

    - Elective and emergency admissions are segregated into zero day stay and non-zero day stay activities
        - National growth in emergency admissions over the past five years is almost exclusively as zero day stays
        - Growth in emergency zero day stays at acute sites is due to step increases. These step increases are due to the opening of various types of assessment & observation units which acts to re-badge A&E attendances as an 'inpatient'
        - Growth in zero-day stay elective admissions also occurs as a step change when otherwise outpatient procedures, tests, injections, non-surgical interventions are re-badged as a 'day case'

- Finally, the model used small area data rather than larger areas such as ward or local authority (LA) boundaries. Larger area models suffer from a lack of specificity due to small pockets of deprivation, age, students or ethnicity being obscured in the average over a larger area. For example, the effect of students cannot be discerned if LA data is used since all LA's tend to cluster too close to the UK average to observe any effect.

- Examples are given for locations and organisations within the South Central Strategic Health Authority. These have been used with permission.

- Opinions & conclusions expressed in this report are those of HCAF.

History of development

The model was initially developed as part of a larger model which predicts the likely change in patient flows under choice, when a site is relocated or when a new competitor site is opened. A large acute trust located at the corner of an SHA was looking to implement radical reconfiguration which was likely to effect patient flows in four SHA's and across 15 to 20 local authority boundaries depending on the various configurations under investigation. Data was only available for six local authority areas and hence a method was needed to predict specialty flows from areas where there was no base data. Actual development to its current form has taken over 18 months of testing and refinement. The model uses lower super output area (LSOA) level data. A LSOA contains around 1,500 head of population and has a high enough level of activity to detect differences in activity in the face of random variation.

An initial version of the model used specialty level data to show the important role for the Index of Multiple Deprivation (IMD) as a factor influencing demand and the fact that elective demand increases far less than emergency demand as IMD increases.

Table One: Percentage increase in activity for a 10 unit increase in IMD.

Specialty	GP-referral to outpatient	EM	EL
Mental Illness	27	-	-
Thoracic Medicine	19	13	7
General Medicine	16	24	4
Rheumatology	14	3	7
Gynaecology	13	4	4
Gastroenterology	12	36	5
ENT	11	10	10
T&O	10	16	6
Medical Group[1]	9	23	3
General Surgery	9	23	4
All Specialties	9	19	4
Paediatrics	7	9	1
Oral Surgery	6	12	3
Urology	5	12	1
Cardiology	5	16	1
Plastic Surgery	3	12	5
Ophthalmology	3	5	7

As can be seen from Table One IMD will play a significant part in the allocation of resources to different aspects of care and explains a significant part of the observed different mix of services and service spends seen in different PCT's.

The role of students was also demonstrated as a factor influencing elective demand. The current generation of the model uses more sophisticated age-adjusted rates specific to student populations and works at HRG chapter level rather than at specialty level.

The concept of an acute site (counting) threshold for the flow of patients from the catchment population was used to explain the different specialty flows at different acute sites. At specialty level this is partly due to specialty overlaps, i.e.

[1] The medical group includes General Medicine, Cardiology, Gastroenterology, Endocrinology, Elderly Medicine, Nephrology, Thoracic Medicine, Infectious Diseases

Gynaecology/Urology/General Surgery have overlaps depending on Consultant interests but it soon became apparent that how patients were allocated to the category 'day case' or 'emergency' admission had profound effects on the reported volumes at different sites.

The need for such a model

The capitation formula distributes a lump sum of money to each PCT. It gives no indication of how this money should be allocated between elective and emergency admission across the different specialties or between specific HRG's or to GP- and other- referred outpatient attendances.

The needs of PBC are far more demanding than a simple calculation of here is your pot of money since each practice needs to know at a small area level how many excess admissions are occurring across a range of HRG's and for GP referral to outpatients.

In addition previous attempts to model demand have suffered from high levels of unexplained variation, i.e. the model parameters and the model structure were only explaining a part of the observed behaviour.

Hence a model was needed which could predict demand at small areas level and that reflected the fundamental factors required to predict UK NHS demand.

Factors leading to the 'unexplained' variation

The approach taken by academics in modelling NHS demand has suffered from a lack of hands on experience resulting in the use of correlations in an attempt to find the parameters 'relevant' to demand. Such an approach may overlook otherwise essential explanatory factors.

Table Two: Thresholds to different types of admission/attendance for musculoskeletal conditions seen in the catchment population surrounding various acute sites[2]

Acute Site	Outpatient first attendance	Elective Admission		Emergency Admission	
		zero day	non-zero day	zero day	non-zero day
Frimley Park	90%	114%	111%	129%	65%
Heatherwood	107%	109%	114%	125%	73%
Banbury	115%	106%	129%	96%	121%
Milton Keynes	118%	82%	92%	126%	103%
Newbury	69%	136%	117%	n/a	n/a
Oxford	107%	94%	82%	99%	105%
Reading	88%	112%	109%	139%	92%
Slough	114%	115%	108%	78%	103%
Stoke Mandeville	108%	106%	97%	55%	94%
Swindon	102%	96%	138%	119%	115%
Wycombe	86%	94%	86%	n/a	103%
Range	**49%**	**54%**	**56%**	**84%**	**56%**

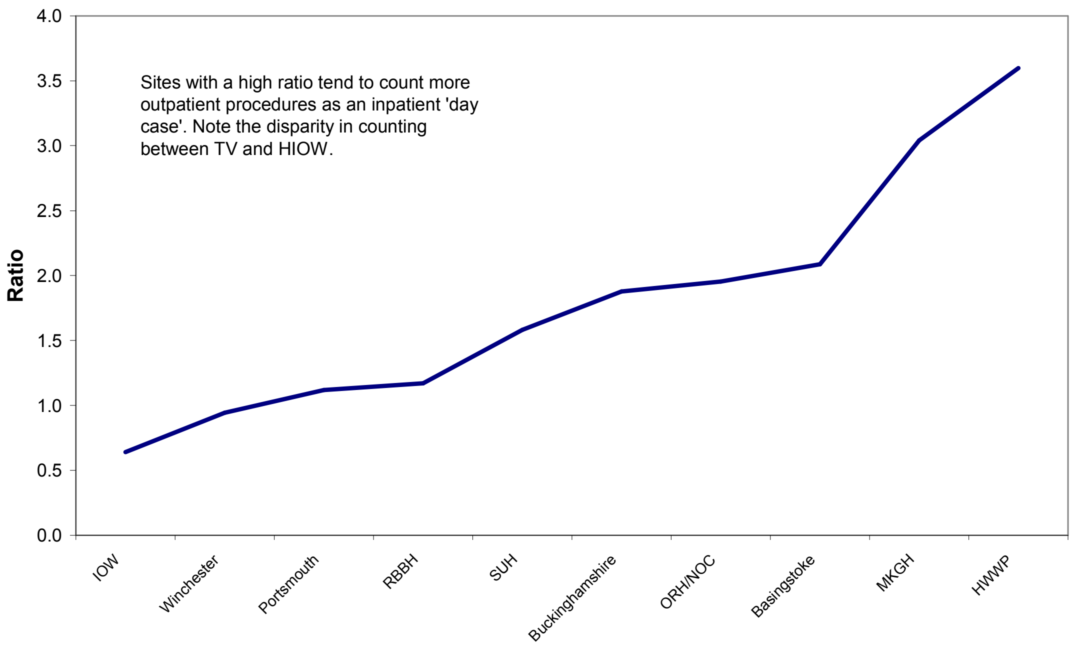
[2] A value of say 136% implies 36% higher admissions/attendance than the overall average. The overall threshold to admission for emergency admission also involves distance specific relationships at certain locations.

This model is based on considerable experience of how healthcare demand behaves in the real world. Experience shows that GP referral behaviour tends to cluster around the acute site to which GP's make most referrals. The system develops its own behaviour based on GP & Consultant relationships. Hence the model assigns each LSOA into a hospital catchment area. Table Two gives examples of the 'threshold' seen for particular acute sites for musculoskeletal conditions.

Additional work was needed to understand the relationships surrounding emergency admission. It was eventually discovered that each acute site had its own unique distance relationship. Hence for Milton Keynes there is a marked increase in emergency admissions for residents living closer to the hospital. While for the Royal Berkshire and Oxford Radcliff hospitals no such distance relationship exists. Other sites have intermediate behaviour. There were no distance related effects for elective demand or for GP referral for first outpatient appointment at any acute site[3].

Likewise students were considered to represent a unique sub-class of the population in that they are generally fit and well in order to be able to successfully study and that they usually have parents coming from more affluent and educated backgrounds. Hence a student living in a particular area will have a healthcare signature which is distinct from the non-student residents. This distinction only appears to apply to elective and outpatient interventions with emergency admissions showing no difference in utilisation.

Figure One: Ratio of non-surgical to genuine surgical day case procedures at different acute trusts. Data is for the top 75 by volume 'day case' procedures at each site.



[3] Other studies have shown that when there is considerable distance involved there may be some decline in expressed elective and GP referral demand. For the residents of Thames Valley the average distance to an acute site is a mere 7 km and the most distant resident (in the western side of Oxfordshire) needs only to travel 29 km to an acute site. Hence the lack of any elective distance relationship is to be expected. That some sites exhibit such strong distance effects for emergency admission is therefore of great interest since it leads to far higher PBR costs.

The model also reflected the basic reality of NHS data in that how a hospital site is structured and works is reflected in how it counts. Since the NHS data dictionary is very vague regarding a number of fundamental and pivotal issues it is not surprising that many HRGs are subject to counting ambiguity, i.e. the same activities are counted as an inpatient at one site but are counted as an A&E attendance, regular day attender or an outpatient attendance at another site. These ambiguities almost invariably manifest themselves in zero-day stays. Hence the model has segregated zero day stay 'inpatient' activities from the non-zero day stay activities.

In combination these simple additions to the model appear to explain the bulk of the so-called 'unexplained' variation seen in previous models.

One of the most significant observations to arise from the model is that it is vital to incorporate these additional factors into any model in order for the model to calculate the correct value of the relationships with IMD, ethnicity, students, etc. Hence this effect will cause the current capitation formula to have incorrect values for the parameters leading to funding misallocation.

Application of the model at LA level

Having developed the model it can then be used to allocate the correct value of each locations share of the national total activity. This calculated share can then be compared to the actual activity and the gap can be converted to money using the PbR tariff.

The steps in this process are as follows:

1. For each HRG the age adjusted share of national average is calculated for each LA using national average admission rates per five year age bands for both elective and emergency admission types. For example, IOW has an age adjusted share of 0.421% of national for HRG H99 (complex elderly with a musculoskeletal diagnosis) but only a 0.206% share for all HRG in Chapter N (maternity & neonatal)
2. Specific additional weighting factors are then calculated using model parameters applied at LA level for both elective and emergency admission. For example, IOW has a day case weighting of 99% for HRG A07 while Winchester has a weighting of 92% for the same HRG.
3. The final share is the product of 1 and 2 above.
4. Benchmarks are then applied for zero day stay activity using 04/05 as the PBR reference point for 2006/07 prices. For example, 47% of emergency admissions for HRG C04 are expected to have a zero day stay while 33% of elective admissions to A01 are expected to be a day case, etc.
5. Excess admissions are calculated.
6. HRG are grouped into those where counting ambiguity is possible and those where it is absent.
7. Performance between different LA is compared and particular LA falling into the catchment area of particular hospitals can be seen to be high cost due to catchment area behaviours. For example, LA in the catchment area of Frimley Park hospital are high cost for those HRG which are subject to counting ambiguity, i.e. this hospital site tends to count more patient contacts as an 'inpatient' than elsewhere.
8. Further analysis can be conducted to see what proportion of the high cost behaviour falls into ambulatory care sensitive admissions, etc.

Outputs from the model

1. HRG where cost savings can be made

Table Three gives the calculated excess cost for the 200 lines leading to the highest excess cost in various locations. It is in these top 200 lines that the potential for greatest cost saving lie. It is important to note that these cost savings will be partly compensated for by a matching set of 200 HRG where there is a large negative excess and that the 'savings' are a mix of counting as well as utilisation issues.

Table Three: Excess cost contained in the top 200 high cost HRG lines for a sample of LA's

Local Authority	Excess Cost			Excess zero day	
	Elective	Emergency	Total	Elective	Emergency
Milton Keynes	£2.5M	£10.5M	£13.0M	2,900	5,600
Basingstoke	£4.2M	£4.6M	£8.8M	4,500	1,100
IOW	£4.0M	£3.0M	£7.0M	2,000	-300
East Hampshire	£1.4M	£5.0M	£6.4M	1,200	350
Slough	£2.0M	£4.0M	£6.0M	2,200	2,600
Cherwell	£3.0M	£3.0M	£6.0M	2,500	1,500
Winchester	£1.3M	£4.6M	£5.9M	1,500	600
Hart	£2.2M	£3.1M	£5.3M	2,600	100
Oxford	£3.0M	£2.2M	£5.2M	1,700	1,000
Reading	£2.0M	£3.0M	£5.0M	2,000	1,700
Total	**£25.6M**	**£43.0M**	**£68.6M**	**18,500**	**12,450**

As can be seen potential cost savings in emergency admission for this sample of LA's are higher than for elective and a significant proportion of the cost savings lie in excess to expected zero day stays, i.e. above national average numbers of outpatient type events re-classified as a 'day case' and above national average volumes of so-called emergency admissions occurring as a zero day stay (in selected locations).

2. Isolating the core problem areas

While Table Three is useful as a means of pointing out total cost savings from a PCT viewpoint it is useful to identify those HRG where the problem is due to excessive admission rates rather than a by-product of counting.

Table Four attempts to make such a comparison by first adjusting for excess zero day stays and then targeting those HRG with a greater than three standard deviation excess of activity. While the adjustment for zero day stays may not be perfect it does allow some interesting core comparison.

This table demonstrates the role of the system behaviour within various acute site catchment areas, i.e. the best locations only needs to tackle some 40 to 60 core HRG while those with higher overall utilisation need to tackle the excess arising from over 100 HRGs.

The fact that Table Four only contains a maximum of 100 HRG while Table Three looked at the 200 HRG with highest excess costs is explained by the fact that from a PCT perspective a significant proportion of potential 'cost' savings are to do with counting issues and that some of the high excess cost lines are with HRG having a less than three standard deviation excess.

Table Four: Total number of HRG where there is a greater than three standard deviation excess of activity after adjusting for the effect of zero day stays.

Main Acute Site	Local Authority	Total	EL	NEL
Portsmouth	Portsmouth	123	82	41
Horton	Cherwell	121	60	61
MKGH	MK	113	12	101
FPH	Rushmoor	110	28	82
Swindon/ORH	VoWH	99	40	59
Basingstoke	Basingstoke	99	22	77
SUH	Southampton	97	55	42
Portsmouth	Havant	95	20	75
SUH/Christchurch	New Forest	93	57	36
FPH	Hart	91	19	72
Portsmouth	Gosport	90	53	37
Winchester/SUH	Test Valley	90	12	78
RBBH/FPH	Wokingham	89	29	60
Portsmouth	Fareham	89	64	25
Basingstoke/Portsmouth	E. Hampshire	88	65	23
RBBH/Swindon	W Berks	87	19	68
ORH	Oxford	80	52	28
IOW	IOW	80	29	51
ORH/Swindon	W. Oxon	76	46	30
Winchester	Winchester	76	8	68
HWWP	WAM	74	18	56
HWWP/FPH	Bracknell	71	22	49
SUH	Eastleigh	63	14	49
RBBH	Reading	59	16	43
HWWP	Slough	58	17	41
RBBH/ORH	S. Oxon	39	22	17

3. A Classification for HRGs

The outputs from the model have also enabled a classification of HRGs according to the level of variation between locations. Some HRG appear to have admission levels which are very similar across very large areas covering many acute sites while others show exceedingly high volumes of ‘admission’ in particular localities.

Table Five gives the results of such a classification and it can be seen that high variation between locations effects over 84% of elective and non-elective activity.

Each HRG has a list of potential counting issues. Other HRG are further classified into ‘ambulatory care sensitive admissions’. The final classification looks at the variation shown by each HRG across a wide range of locations. Some HRG show little variation between locations while others show extreme (>20 standard deviation) variation.

HRG showing extreme variation arise when one acute site counts events in a way that is totally different to national norms. For example MKGH uses HRG S21 to count what may otherwise be considered as regular day attender’s as a ‘day case’ while Southampton and Portsmouth hospitals appear to use HRG E03 in a similar way.

Hence each PCT can use the three classifications to query the counting or coding of events at their local acute site or in the case of ambulatory care sensitive conditions to investigate non-acute ways of meeting the needs of particular patient groups.

Table Five: A classification of HRG based on variation between locations

Variation between locations	Type	% of HRG's	% of Activity	Comments
High	Data Definitions & Counting	10%	37%	Mainly endoscopic and other procedures, tests, injections where the boundary between outpatient and day case is not rigorously applied or regular day attender's incorrectly counted as day case.
High	Clinical Pathways & Intervention Rates	24%	47%	Almost all elective Orthopaedic interventions fall into this category, i.e. there is considerable discretion regarding the decision to intervene or admit.
Low	Data Definitions & Counting	4%	2%	Includes a minority of HRG where there is the potential for isolated instances of ambiguity.
Low	No issues	60%	13%	This is the common core of HRG which shows very little variation between locations & hospitals. This group is almost exclusively low volume.

4. Achieving the savings

It would appear that a significant part of the excess cost experienced by PCT's is due to local acute Trusts counting in a way which is significantly different to what may be considered as 'national average'.

This has considerable unintended PBR cost implications which need to be resolved. The suggested approach is for PCT's to pay at full tariff up to what would otherwise be the national average zero day volume expected for each HRG. Any activity above this level should be costed at the appropriate A&E, outpatient or regular day attender price.

This approach would remove the bulk of the excess zero day stay elective cost experienced by most PCT's as seen in Table Two. It goes without saying that genuine day case surgical interventions are excluded from such considerations.

For emergency admissions it is clear that while excess zero day stays may account for up to £10 M of the excess cost seen in Table Two this still leaves some £33M of potential excess cost due to high levels of emergency admission in specific locations.

Hence the model demonstrates that the bulk of cost savings lie in the more difficult to achieve reductions in overall levels of emergency admission which will rely upon a reduction in the level of primary care sensitive emergency admissions and other HRG's where particular acute sites have a lower than average threshold to admission.