Plain Explanation or Special Pleading?

By Mervyn Stone (January 2013)

Online Report: January 2013

55 Tufton Street
London SW1P 3QL

Civitas is a registered charity (no. 1085494) and a company limited by guarantee, registered in
England and Wales (no. 04023541)

email: info@civitas.org.uk

Independence: Civitas: Institute for the Study of Civil Society is a registered educational charity (No.1085494) and a company limited by guarantee (No. 04023541). Civitas is financed from a variety of private sources to avoid over-reliance on any single or small group of donors. All the Institute's publications seek to further its objective of promoting the advancement of learning.

Negative developments

Things have moved on since I wrote 'Doctors in the Waiting Room' for Civitas. The Minister's door has opened and GPs now know that the NHS Commissioning Board has put innovation on the back-burner for 2013-14, pending an 'urgent, fundamental review'[1] of what to do for the 2014-15 allocations. The innovation would have been the replacement of a PCT-funding formula (getting rather too long in the tooth and difficult to defend) by a new sort of formula with the acronym PBRA for Person-Based Resource Allocation. It is a formula for funding the bulk of hospital inpatient & outpatient care—one that the Nuffield Trust has worked on for years since The House of Commons Health Committee recommended the direct measurement of care in 2006.

In December 2011, the Trust released a report[2] by its Head of Research, Martin Bardsley, and Director, Jennifer Dixon, alluding to a 'development' of the PBRA model[3] that the British Medical Journal had published in November 2011 (and that my Civitas article took to pieces in order to see how it might work). The report was proud of what the development had achieved and was aiming at—that it would soon be put to work:

Subsequent development has increased ... predictive power to 85% [from 77% for PBRA] ... [In order to] update the formula to set target allocations in 2012/13 more sophisticated ways of incorporating diagnostic information were tested and recommended, for example taking into account the possible impact on costs of having certain combinations of diagnoses.[2]

My curiosity about these developments was politely rebuffed when Dr Dixon informed me (on New Year's Day 2013) that the Trust's website contained all the material on PBRA that had by then been released. One week later, in happy conjunction, the Department of Health published a by-then 15 months old 'update'[4] on those developments to its advisory committee on resource allocation (ACRA). This update of earlier PBRA models was a draft report, by a new-titled team of health economists led by Dr Dixon, on a project apparently started in 2010 (not 2011 as the report inconsistently states).

The model PBRA3 recommended in the update makes huge changes to the PBRA model 3. The PBRA3 Team does not defend the financially consequential changes that have been made between the models—both of which have been strongly recommended by much the same group of researchers. These appear to be sidelined in the drive to get the overall 'predictive power' of the model as high as possible. PBRA3 is a composite of three sub-models for the under-15s, the over 65s and the rest of us. The fact that the sub-models were, in effect, independently fitted will have contributed to the increase in the 'predictive power' (assessed by the relevant R^2 as from 77% to 85%), stemming from the model-enriching interactions with age that make the model much more flexible and adaptive to what the data is trying to communicate to the modellers. PBRA3 also greatly reduced from 5% to 3% the proportion of statistically significant negative coefficients (like the minus £436 coefficient for dementia defined and reported elsewhere[5]). The PBRA3 Team has three explanations for accepting the ones that remain.

Of the morbidity flags 88% of them had coefficients that were positive and statistically significant, indicating a tendency to increase costs and were statistically significant (see table 9 for a selection). A small proportion (around 3%) of the morbidity flags were negative i.e. associated with lower future costs in hospital care. These can be explained either as ***conditions that are associated with death****; some* ***conditions whereby treatment reduces the likelihood of further problems e.g. removal*** *[of]* ***appendix,*** *and* ***some infectious diseases or conditions where costs from the dependent variable may have been excluded, mental health, specialist care.***[4]

For another negative coefficient (not a morbidity flag) the report finds an explanation in ACRA's rich collection of concepts—an 'unmet need' revealed by an expectation-offending 'wrong sign', now called a 'differentially met need'.

The black and minority ethnic group proportion has a negative coefficient for age group 15-64. We treat this as an indication of differentially met need and do not include its effect when calculating needs-based predictions at practice level as it would lead to lower (all else equal) allocations for practices with a higher proportion of ethnic minorities. [4]

This paper takes a close look at the logic or otherwise of these quotations, after tabulating a selection of the relevant coefficients in PBRA and PBRA3 and presenting a simple (precise & interpretable) mathematical expression for the coefficients of 'morbidity flags' (elsewhere[5] identified as 150 dummy, 'whether-or-not', variables of ICD10 diagnoses).

Tabulations and clarifications

Table 1: Statistically significant coefficients in PBRA [5]

F00-F03	−£436	Dementia
K35-K38	−£93	Appendix
N40-N51	−£93	Male genetic organs
N99	−£349	Genitourinary
O10-O75…	−£34	Labour & delivery
S70-S79	−£72	Hip & thigh
Z30-Z39	−£86	Reproduction

Table 2: Some statistically significant coefficients in PBRA3 [4]

ICD code	Age 0-14	Age 15-64	Over 65	Disease
A80-A89	−£558	−£129	£606	Viral CNS
A90-A99	−£97	−£529	−£639	Arthropod-borne viral
B20-B24	−£1639	−£667	−£942	HIV
B50-B64	£155	−£352	−£1462	Protozoal infection
C00-C14	−£3472	£1016	£1104	Malignant neoplasm
F00-F03	£10871	−£52	−£250	Dementia
F04-F09	£5997	−£26	£244	Other organic mental
K65-K67	−£1712	£636	£396	Peritoneal
P05-P96	£383	−£172	−£1143	Perinatal

How should we understand and interpret these numbers? The PBRA model may be obsolete but it is at least fully documented [2, 3, 6, 7] —so that is where we will try to answer it. At one level, treating the least-squares model-fitting process as a black box, there is no additional need for understanding. At that level, all we need to know is that the PBRA team fed their computer with a huge data-base of about 5 million individual records on over 300 predictor variables (for individuals on GP registers on April 1 2007) and that, from the numbers that emerged, a prediction could be made of how much it would cost an individual's GP to commission hospital care for that individual in 2007-08. The prediction is the evaluation of a formula that would have credited the GP's account with the PCT (or CCG in the future) for every item in the individual's record that might require hospital treatment. For example, the first coefficient in Table 1, £436, is what the PBRA formula would have debited from the 2007-08 PCT account of a practice for looking after a 70-year-old who had had a 'completed consultant episode' at some hospital in 2006-07 recorded as 'dementia' in the Hospital Episode Statistics database.

At the deeper level of understanding that will be needed if we are to explain the negative coefficients that debit the GP's account, we need to look at PBRA's 200-year-old statistical methodology in action (able to do things with speeding electrons that its inventor, Karl Friedrich Gauss, would not have thought sensible in such a theoretically vacuous subject-matter). We have to open the black box, rerun the least-squares machinery at an infinitesimally slow speed to imagine that we can see exactly how, for example, the £436 for dementia in Table 1 from billions of double-precision arithmetical operations of the statistical-computing software STATA version 10.

What follows is a hypothetical description of the final stage of a computation that would almost certainly differ in shape from any stage of the STATA computation, but which would give the same number £436 (if we can ignore cumulative rounding errors in the billions of STATA operations, some of which may be required to subtract a very big number from an almost equally big one, reflecting high correlations between the inputted predictor variables). I will make the immaterial but exploitable assumption that the ICD code for dementia was the last variable that the computation had to deal with since the order in which the—immaterial because the order in which variables are processed does not affect the final outcome.

Before the least-squares fitting is hypothetically completed by the dementia variable, each of the 5 million individuals will have both a penultimate prediction of actual hospital cost (waiting to be completed) and an associated residual or *exceedance* defined as *actual cost minus penultimate prediction*. (In general, a penultimate prediction will not be the same as the predictor that is left when the final 'adjustment' for dementia in the 'full formula' is 'frozen'—the jargon in quotes here refers to the alternative but equivalent formulation of the model as a sequence of up-and-down adjustments[5] around the national average cost.) The least-squares fitting ensures that average of these 5m exceedances will be zero. But we cannot have same expectation for the average of those individuals who have a 'morbidity flag' (i.e. a 1 rather than a 0) for the binary dummy variable for dementia—the dementia cases defined as those who had a hospital episode recorded as dementia in the 2005-07 data-base. To say more than that, we need just one technical concept:

The square R^2 of the-data base correlation coefficient between the dementia dummy and the weighted sum (linear combination) of the 349 other variables in PBRA with which it is most highly correlated.

The theory of least-squares shows that statistic R^2 is needed to express the dementia coefficient as the simple ratio:

***Dementia coefficient* = (*Average exceedance for dementia cases*) /(1 − R^2)**

The same expression holds for any of the dummy variables in the model, not just the ICD10 codes, and statisticians will recognize R^2 as the easily calculated 'multiple correlation coefficient' between the dementia dummy and the other variables. Whether or not R^2 is large enough to be an important factor in interpreting the minus £436 could be resolved by retrospective computation of the PBRA formula data-base, if it has been preserved. We know that most dementia cases are elderly. But most of the population is not elderly and most of the elderly do not have dementia, so correlation with age would not be big enough on its own to make the $1-R^2$ influential in the interpretation of the minus £436. But it could do so if there were another diagnosis correlating highly with dementia—in other words, if an appreciable proportion of the cases of each of the two diagnoses were for individuals who had had both so that, for example, $1-R^2$ would be about 50% for a second diagnosis with the same incidence as dementia and a shared proportion of cases of about 70%.

In what follows, we do not need to make any conjectures about the size of R^2 in the above formula for an ICD10 coefficient. It suffices that the coefficient has the same sign as the numerator. Argument will be based on the precise character and provenance of the numerator—and, specifically, on the fact that a statistically significant coefficient implies a statistically significant numerator of the same sign.

Completing the logic

The statistically significant minus for the dementia coefficient therefore becomes a statistically significant minus for the average penultimate exceedance (cost minus prediction) for the dementia cases in the data-base. If the average had been positive rather than negative, that would not, in itself, be a reason for questioning the validity of the model—the fitting would have been completed to give a positive estimate of the dementia coefficient. As it is, the PBRA researchers ought to accept the fact that the highly statistically-significant negative is compelling evidence that their model is not 'true' i.e. it is not an adequate description of the 'real' (operationally acceptable) relationship with actual cost that could be used for prediction without further refinement. Sceptics about the whole approach have no difficulty with that and will therefore have no problem accepting the following possible explanation of the negative dementia coefficient.

Dementia cases are on the whole elderly—even more aged than the 'elderly'. I will risk just one conjecture: that the negative average for the dementia cases is a consequence of a negative for a larger class (the aged) in which they are, on the whole, to be found—which means either that the average formula prediction or this class is unrealistically too high (reflecting some inadequacy) or that the average actual cost is too low (perhaps reflecting the painful history of resource allocation).

The second quotation of the PBRA3 update would appear to attribute the dementia negatives for F00-F03 in Table 2 to the omission of mental health costs from the dependent variable of actual cost. The F00-F03 Dementia code comes under 'Mental and behavioral disorders' in the WHO ICD10 codes listing. The BMJ paper [3] confirms that costs for associated hospital episodes were indeed excluded from the dependent variable for PBRA. But if dementia cases do not come into PBRA's actual cost, how do dementia dummies come to be fitted, as they are, with highly significant coefficients (of any sign) if there is no trace of the cost of dementia cases in the dependent variable—*if the model is true*? In other words, how are the dementia cases identified at the (hypothesized) penultimate stage of the fitting if they do not participate in the costing? Without some form of magic, the findings must be attributed to empirical inadequacies of the model—and to a dementia-dummy correlation with age that picks up the gap, for the aged, between actual and predicted cost of hundreds of pounds (assuming that R^2 is not responsible for inflating a small gap to the level of hundreds).

The same clash of explanations applies to the −£52 and −£250 in Table 2. The −£250 for the over-65s is consistent with the argument for PBRA, whereas the −£52 for those aged 15-64 may be coming mainly from the over 50s where different factors influence either the average prediction or the average cost. The eye-catchingly positive figure of £10,871 for children surely requires explanation by the PBRA3 Team. Could it be due to massive hospital costs for non-dementia diagnoses that such children may have to suffer and that are not excluded from the actual hospital costing as mental health disorders'—or could it be that elusive possibility, a concomitant diagnosis (ICD10 dummy) that correlates so strongly with the dementia dummy that $1-R^2$ is close enough to zero to inflate a merely positive numerator to the thousands level?

Out for the count!

It remains to consider the strikingly regular morbidity count coefficients in PBRA3 in Table 3—all of them highly statistically significant. Their remarkable smoothness and regularity contrasts with the haphazard differences between the coefficients in the sub-models of Table 2. Did the PBRA3 Team not question the logic of this aspect of their modelling once they had tabulated the results?

Table 3: A redacted version of the update's Table 8 [4].

Morbidity count	Age 0-14	Age 15-64	Age over 65
2	−£230	−£238	−£9
3	−£384	−£409	−£98
4	−£514	−£566	−£186
5	−£624	−£666	−£269
6+	−£653	−£863	−£311

The numbers in the table are the coefficients for 5 dummy variables that identify whether or not an individual had a specified number of ICD10 diagnoses (from 2 to 6+) in the two year period 2007-09. The update has a succinct and superficially plausible explanation of these figures:

The impact of the variables that count different morbidities are negative (see Table 8) indicating that the additional costs of having more than one condition is less than the sum of cost of the two individual components.

While it is manifestly true that the negatives have an 'impact', it is only an impact *on the value of the formula.* The PBRA3 Team's explanation veers from reality (in real time!) in its claim for what that is indicating. Their explanation rests on an implicit trust that the formula was a reasonable approximation in 2009-10 to the average actual cost in large enough groups of individuals such as GP practices—in other words that the model was true in that broad sense.

To see where the PBRA3 Team's 'indicating' goes seriously wrong, simply focus on one of the coefficients in the table e.g. the −£653 for the group of children who had at least six morbidities in total over the two years 2007-09 (G6+ for short). The £653 is the average for the G6+ children of the amount by which their predictions from their 2007-09 record exceed their actual costs in 2009-10—inflated (via the $1\text{-}R^2$ denominator) by whatever correlation there is between the dummy for a count of 6+ and the other variables (necessarily more than the correlation with the sum of the ICD10 dummies). It does not put much stress on intuition to suggest that this average is positive (as it turned out to be). Unless there had been a dramatic increase in childhood diseases between 2007-09 and 2009-10, the G6+ actual cost average would come from an on-average smaller set of diagnoses even if we were comparing like with like. The natural phenomenon of 'regression to the mean' may be expected to apply here, just as it does in other areas of life—there has to be a rebalancing from year to year as some of those with under six diagnoses in overlapping two-year periods move up to the six plus level and some of those at that level move down. The reduction is compounded by the fact that, as far as the assessment of the coefficient is concerned, we are comparing the diagnoses in one year with those over a two-year period.

That the coefficient should so smoothly decrease in magnitude, as we go down from a count of 6+ (in the data-base) to a count of 2, should also not stress any intuition—the chance-determined reduction in disease incidence would be expected to be less for the lower count groups. Even if the reduction had reversed sign for a count below some value, our explanation of the uniform negatives can again invoke the compounding effect of halving of the time period between construction of the predictor and its application.

Conclusion and recommendation

Rather than repeating the now well-known objections of leading statisticians to how the Department of Health ignores statistical principles and logic in constructing its resource allocation formulae, this paper has exposed the logic of claims in a report [4] on the latest offering to the Department—claims that smack of special pleading. Our negative findings lend strong support to any recommendation that a clean sweep of the current predictive approach should be the final step in the Commissioning Board's 'urgent, fundamental review'.[1] The first step could be withdrawal of its almost unqualified blessing of the PBRA3 formula:

… the formula proposed by the Advisory Committee on Resource Allocation accurately predicts the future spending requirements of CCGs [1].

---

1 'Everyone Counts: Planning for Patients 2013/14' (2012) NHS Commissioning Board, Everyone Counts: Planning, Bulletin for proposed CCGs: Issue 25, 18 December. At: http://www.commissioningboard.nhs.uk/2012/12/17/everyonecounts/#comment-28278 [accessed 30 January 2013]

2 Bardsley, M. and Dixon, J., 'Person-based Resource Allocation: New approaches to estimating commissioning budgets for GP practices' (2011) The Nuffield Trust, December 2011. At: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/person_based_resource_summary_dec11_0.pdf [accessed 30 January 2013]

3 Dixon, J., Smith, P., Gravelle, H., Martin, S., Bardsley, M., Rice, N., Georgiou, T., Dusheiko, M., Billings, J., De Lorenzo, M., & Sanderson, C., 'A person based formula for allocating commissioning funds to general practices in England: development of a statistical model' (2011) *British Medical Journal*, 343, 22 November. At: http://www.bmj.com/content/343/bmj.d6608 [accessed 30 January 2013]

4 PBRA3 Team, 'Developing a PBRA formula for allocations in 2012/13: Update for ACRA' (2011) The Nuffield Trust, Evidence for Better Health Care, 22 September.
At: https://www.wp.dh.gov.uk/publications/files/2013/01/ACRA201119-GP-Commissioning-General-and-Acute-PBRA-update.pdf [accessed 30 January 2013]

5 Dixon, J., Smith, P., Gravelle, H., Martin, S., Bardsley, M., Rice, N., Georghiou, T., Dusheiko, M., Billings, J., Lorenzo, M.D. & Sanderson, C, 'A person based formula for allocating commissioning funds to general practices in England: development of a statistical model' (2011) *British Medical Journal*, 343, 22 November. London School of Hygiene & Tropical Medicine.
At: http://www.researchonline.lshtm.ac.uk/30430/1/dixj872671 [accessed 30 January 2013]

6 Stone, M., 'Rejecting anempirical 'person-based' formula for funding CCGs in favour of the farming analogue of one-year ahead extrapolation' (2012) In the Civitas blog, 'Doctors in the Waiting Room', 4 December. At: http://civitas.org.uk/newblog/doctors-in-the-waiting-room/
[accessed 30 January 2013]

7 PBRA Team, 'Developing a person based resource allocation formula for allocations to general practices in England' (2009) The Nuffield Trust, 14 October. At: http://www.nuffieldtrust.org.uk/sites/files/nuffield/document/Developing_a_person-based_resource_allocation_formula_REPORT.pdf [accessed 30 January 2013]