Multilevel assessment of immunisation uptake as a performance measure in general practice

Kelvyn Jones, Graham Moon

Abstract

Objective—To provide a measure of general practitioners' performance regarding uptake of immunisation against pertussis, taking account of the impact of patient characteristics on levels of uptake.

Design—Multilevel model of immunisation status against six measures of patient characteristics (level 1 predictor variables) with practice constraints as level 2 variables.

Setting—126 practices in southern England.

Subjects—2048 infants identified from infant surveillance and immunisation records.

Main outcome measures—Top 10 practices with respect to uptake of pertussis immunisation according to a "null" model (model A) and according to a model that included six level 1 variables (model B); differences in ranking between the two models.

Results—Practices with low numbers of infants requiring immunisation had imprecise and unstable uptake rates (range 0%-100%). With the multilevel procedure, after controlling for patient characteristics, practices in suburban catchment areas comprised largely of mature or young professionals performed best. Most improved performances when patient characteristics were taken into account were in practices in areas with a stable population and local authority housing—one such practice improved its ranking by 47 places.

Conclusions—Crude uptake rates are inadequate performance indicators. Alternative approaches suggest that praiseworthy efforts to raise immunisation rates in unpromising areas are unrewarded by simple target based assessments.

Introduction

The recent NHS reforms have brought about fundamental changes that draw extensively on management techniques from the private sector.[1,2] Central to the reformed service are enhanced systems for performance review. By citing comparative performance, these are intended to indicate to clinicians and managers where improvements to services might be made.[3-6] The output from these systems is generally quantitative, but is often taken to imply quality. More importantly, these quantitative output indicators are being linked with future resourcing and, in the case of general practice, to schemes for target payments.[7] As reputation and reward will come to depend on quantitative performance, it is vital that valid and reliable indicators are used.

Although increasingly pervasive in hospital and community health services, performance indicators in general practice have, to date, been uncommon. With the advent of the new general practitioner contract this is set to change.[7] Requirements for annual reports from practices will generate de facto performance indicators, enabling comparison among practices.[8] Furthermore, for childhood immunisation and cervical cytology fee for service payments have been replaced by payments linked to target uptakes.[9]

In this paper we focus on the shortcomings of performance indicators by studying childhood immunisation uptake as a performance indicator in general practice. It has been recommended that practice annual reports should include reference to this subject, and there is some evidence that the publication of immunisation rates of practices engenders competition among practices to improve uptake.[10,11] Moreover, practices are now to be paid according to two bands of uptake: those practices achieving the World Health Organisation's immunisation target of a 90% uptake will receive a higher rate of pay and those exceeding 70% uptake a lower rate. Performance is thus reduced to crude percentage uptake thresholds.

Methods

PROBLEMS OF CRUDE RATES

Performance indicators are generally expressed as crude aggregate rates. For immunisation these are the numbers of children receiving an immunisation as a percentage of the total target population of children aged under 2 years for the year of study.[7] These rates may be used to generate crude league tables. Such league tables often hide more than they disclose because crude rates conflate several distinct sources of variation. Firstly, there is an overall performance, irrespective of clients' characteristics. Secondly, uptake may reflect the composition of the practice list—housing tenure, for example, has an effect on uptake; practices comprising mainly people who live in local authority housing can be expected to have a low uptake rate.[12] Thirdly, the crude rate obscures the possibility that practices do not perform equally well for all patients. A reasonable rate may be achieved through providing an excellent service for some patients and an abominable one for others.

These factors limit the usefulness of crude rates. They have been intuitively recognised by general practitioners; there are several published examples of attempts to change the composition of a list by excluding problem housing estates or problem groups.[13,14] Before a valid peformance indicator can be generated methods are needed which allow patients' characteristics to be taken into account while simultaneously recognising that uptake by clients may be affected by the characteristics of particular service providers. Similar problems have been recognised in research on education, where studies have been focused on the relative importance of pupils' characteristics and their school's resources in determining the overall performance of a school.[15,16]

MULTILEVEL MODELLING

Multilevel modelling[17,18] can be applied to these

Department of Geography, Portsmouth Polytechnic, Portsmouth, Hampshire PO1 3HE
Kelvyn Jones, PHD, *senior lecturer in medical geography*

School of Social and Historical Studies, Portsmouth Polytechnic, Portsmouth, Hampshire PO1 3AS
Graham Moon, PHD, *senior lecturer in health studies*

Correspondence to:
Dr Moon.

BMJ 1991;303:28-31

requirements and has considerable potential in the analysis of primary health care statistics.[19,20] It can be conceived as an extension of the generalised linear model[21] in which the individual decision to immunise is explained by individual patient (level 1) and practice (level 2) constraints. The decision to immunise is thus the response variable, and it is related simultaneously to both level 1 and level 2 variables by using a logit link function and a binomial error term; this is the appropriate form for a model with a binary response variable.[22] The alternative would be to conduct an ecological analysis whereby client characteristics are represented by an aggregate variable for each practice and the modelling operates exclusively at level 2. This would be inappropriate and invalid, as shown in the published work on education.[16]

Multilevel models, like any model, are generalised representations of reality. It is unlikely that all potentially influential factors at level 1 and level 2 will be included; a measure of variance in the response variable will be unaccounted for. Multilevel models allow this residual variance to be apportioned to the appropriate level. If all relevant patient characteristics are included in the model as level 1 variables, thereby maximising the variance explained at the individual level, the unexplained variance remaining at level 2 will represent an improved, contextualised measure of practice performance. A further refinement, not undertaken in this paper, would be to include level 2 variables, such as practice size or a measure of professional knowledge concerning contraindications, to account for differential performance.[20] In this paper we focus on patient mix as a potential reason for performance variation.

DATA AND ANALYSIS

The utility of the multilevel modelling approach in improving performance indicators was evaluated by using a two level dataset. This comprised 2048 children aged under 2 years in 126 practices. The response variable categorised the children by their pertussis immune status (all should have had completed schedules). Six level 1 variables were selected, reflecting factors thought to constrain parental decisions regarding immunisation (table I).[23] The practices constituted the level 2 grouping. The full dataset was produced by linking an immunisation call and recall database with infant surveillance data.

TABLE I—*Level 1 variables thought to be concerned with constraint of parents regarding immunisation against pertussis*

Variable	Hypothesis
Previous death of an infant	Protectiveness; positive attitude to preventive medicine; link with high uptake
Mother smokes	Negative attitude to health promotion; associated with low uptake
Housing tenure	Rented tenures associated with low uptake
Stability of family	Unstable family relations may place undue burden on a single parent to ensure attendance for immunisation
Employment status	Surrogate for social status; unemployment likely to be associated with low uptake
Mother's age	Older mothers more likely to challenge professional recommendation to immunise

Multilevel estimation was undertaken with a VAX8350 by variance components analysis (VARCL), a modelling package allowing choice of link function and error structure.[24]

Two models of uptake of pertussis immunisation were estimated. Model A, the "null" model, identified the general mean uptake in the entire dataset and apportioned variation to level 1 or level 2. Finding no variation at level 2 would imply that although patients might vary in their immunisation status, there was no variation among clinics. Conversely, confirmation of variation at level 2 would indicate either genuine differences among practices or compositional effects due to systematic differences between practices in terms of client characteristics. Model B was designed to evaluate the possibility of compositional effects by including the six level 1 measures and thus effectively controlling for patient characteristics. A significant level 2 variance in model B would indicate that, given the overall mean uptake and client characteristics, meaningful variations in uptake among practices remained.

Level 2 residual variance for each practice can be calculated from the overall level 2 variance. In model B these residual variances may be taken as measures of the performance of a practice, given the overall level of uptake across all practices and the specific mix of client characteristics in each practice. Clearly the choice of level 1 variables, the specification of the model, can affect the size and sign of the residual variance, but, within this limitation, the level 2 residual variance provides the required improved, contextualised performance measure and can be ranked to construct league tables indicating comparative performance.

Results

A traditional analysis of practice performance regarding immunisation uptake would be based on dividing the number of immunised children by the target population for each practice, calculating percentage uptake rates, and making comparisons among practices. Table II sets out these aggregate results, which indicate that there is a wide spread of performance. Indeed, while 12 of 126 practices achieved uptakes of 100%, four had an uptake of 0%. Overall these results suggest that many practices have some way to go before achieving a 90% uptake, at least for pertussis immunisation.

As already discussed, these crude rates are poor measures of performance because they disregard client characteristics. They may also be unreliable estimates of true performance if there are few children in the target population. Those practices with few children are more susceptible to being classed as having poor or good performance. To take an extreme example, failure to immunise one child in a practice target population of one can have a catastrophic impact on an uptake rate. At the heart of this problem of small numbers is the separate estimation of each practice uptake. The multilevel procedure operates more effectively by considering small practices in relation to the overall performance of all practices. Rates for practices with unreliably small target numbers are shrunk towards the overall performance. Reliable rates are unaffected by this process. Considerable theoretical and practical research has indicated that this procedure leads to much improved estimation of true long term performance.[25]

The effects of this shrinkage procedure are most evident when practices ranked by the percentage uptake indicator are compared with a similar ordering of practices according to the level 2 residuals in model A. Only two practices appeared in the top 10 best performers on both percentage uptake and by model A. They had target populations of 10 and managed to achieve 100% uptake. Practices with lower target populations were down ranked. The multilevel procedure thus recognises achievement, but values it in proportion to the size of the target population. Consequently, in model A, the best performing practice becomes one which has successfully immunised 95% of its relatively large target population of 31. Table III sets out the 10 best practices by model A together with their uptake rates and target numbers. They were

TABLE II—*Crude percentage uptake of pertussis vaccine as measure of practice performance (n=126)*

Uptake (%)	No of practices
100	12
90-99	2
70-89	48
1-69	60
0	4

TABLE III—*Top 10 practices based on results of uptake of pertussis immunisation by model A*

Position	Practice No	Uptake (%)	Target population	Type of patients/location
1	76	96·7	3[illegible]	Young professionals; commuter village
2	13	100	18	Mature professionals; suburb
3	80	85·2	54	Young professionals; suburb
4	86	100	11	Stable local authority estate
5	98	84·8	33	Mature professionals; suburb
6	53	81·7	60	Inner city improvement area
7	93	86·9	23	Mature professionals; resort
8	92	60·0	75	Working class people; urban
9	48	100	9	Mature professionals; suburb
10	55	84·6	26	Middle class people; urban

typically located in suburban areas characterised by middle class professional households.

Model B takes account of the constraints which patient characteristics place on uptake rates. The results showed significant relations ($p=0·05$) between uptake and tenure, with local authority and private renting tenants being less likely to have their children immunised. They also, surprisingly, suggested that mothers who smoked were more likely to have their child immunised. The effects of the other predictor variables were not significant. Use of the model in a predictive sense indicated an uptake of over 70% for employed owner occupiers but only 45% for unemployed local authority tenants.

Differences in practice performance were still evident in model B. The best performances when patient characteristics were taken into account were still those of practices serving middle class suburban professionals. Table IV gives the 10 best practices by model B. With two exceptions these were the same

TABLE IV—*Top 10 practices based on results of uptake of pertussis immunisation by model B*

Position	Practice No	Uptake (%)	Target population	Type of patients/location
1	76	96·7	31	Young professionals; commuter village
2	13	100	18	Mature professionals; suburb
3	80	85·2	54	Young professionals; suburb
4	92	60·2	75	Working class people; urban
5	53	81·7	60	Inner city improvement area
6	86	100	11	Stable local authority estate
7	98	84·8	33	Mature professionals; suburb
8	54	79·0	43	Inner city; mixed tenures
9	48	100	9	Mature professionals; suburb
10	77	96·7	31	Young professionals; mixed tenures

practices as with model A, although some reordering within the top 10 occurred. More interestingly, however, table V shows those practices whose ranking by model A changed by over 20 places when analysed by model B. Those whose performance improved by 20 places were primarily large practices based in health centres and servicing local authority estates with stable childrearing populations in both inner city and peripheral locations. Worse performances characterised those practices in areas with older populations of owner occupiers. These results suggest that practices where childhood immunisation is a recurring feature of the workload may be making strenuous efforts to overcome any equation between local authority tenure and lower immunisation rates.

Discussion

The Department of Health Performance Indicators Group has urged that performance should be considered in context.[26] Multilevel modelling with data on patients within practices provides an appropriate unified approach for producing performance indicators in general practice. We acknowledge the shortcomings of the analysis in this paper. The selected level 1 variables, despite empirical justification, could be challenged for their adequacy in capturing patient characteristics fully, although they were shown to have an effect on uptake. Furthermore, data on patients are required for a valid analysis but may be subject to restricted access. There may be particular level 2 factors, such as membership of a computerised call and recall system or being a single handed general practitioner, which might inhibit performance and excuse poor uptake. The robustness of the relatively new technique of multilevel modelling with regard to factors such as multicollinearity are, as yet, not fully understood. Furthermore, immunisation against pertussis is a specific and perhaps problematic case given the controversy over sequelae; work is in progress to assess the extent to which other childhood immunisations conform to the same pattern.

Notwithstanding these caveats, the two models reported here indicate three important conclusions. Firstly, performances based on small target populations are inherently unstable. An achieved performance should be considered in the context of the size of the target population. The effort required to immunise successfully a large target population is proportionately greater than that required in practices with very few infants.

Secondly, the social context provided by patient characteristics undoubtedly affects the achieved level of performance. Given a list dominated by patients with characteristics known to be associated with low uptake, some practices "perform" very well yet have low uptakes. A rational, contextualised peformance indicator should acknowledge this; indeed, those practices that are particularly effective with low uptake groups may provide models of good practice.

Thirdly, and perhaps most importantly, performance targets should not be seen as absolute. They should be considered in relation to the social context in which they are or are not attained. Among the top 10 performing practices by model B (table IV) were five practices that did not attain the World Health Organisation's target of 90%. None of the practices achieving substantially improved performances when their list characteristics were taken into account (table V) had uptake levels above 75%. The conclusion must

TABLE V—*Practices whose ranking by model A changed by more than 20 places with model B*

Practice No	Uptake (%)	Change in ranking	Type of patients/location
115	66·7	+47	Stable peripheral local authority estate
21*	53·8	+35	Stable peripheral local authority estate
107	50·0	+34	Stable peripheral local authority estate
30*	64·8	+28	Mixed tenures; suburb
105	57·1	+23	Stable peripheral local authority estate
16*	46·1	+23	Stable peripheral local authority estate
106	75·0	+22	Stable inner city local authority estate
70*	71·4	+21	Stable inner city local authority estate
29*	59·5	+21	Mixed tenures; suburb
17*	66·6	+20	Stable peripheral local authority estate
104*	66·7	−24	Inner city working class people; elderly people
65	66·7	−25	Middle class elderly people; urban
102*	66·6	−32	Mature professionals; urban
37	75·6	−36	Mature professionals and elderly people; suburb

*Practices based health centres.

be that some practices perform excellently but still do not reach performance targets. The implication is that these practices will, despite their best efforts, not qualify for target payments. The fear is that without some acknowledgement for their effort they will either artificially manipulate their list to exclude "problem" patients or they will stop trying to attain high uptake.

1 Secretaries of State for Health, Wales, Northern Ireland, and Scotland. *Working for patients*. London: HMSO, 1989. (Cmnd 555.)
2 Department of Health and Social Security. *NHS management inquiry*. London: HMSO, 1983. (Griffiths report.)
3 James M, Roberts A. Performance indicators for the National Health Service. *Health Trends* 1987;**19**:12-3.
4 Lowry S. Focus on performance indicators. *BMJ* 1988;**296**:992-4.

5 Skinner P, Riley D, Thomas E. Use and abuse of performance indicators. *BMJ* 1988;**297**:1256-9.
6 Allen D, Harley M, Makinson G. Performance indicators in the National Health Service. *Social Policy and Administration* 1987;**21**:70-84.
7 Health Departments of Great Britain. *General practice in the National Health Service: the 1990 contract.* London: Health Departments of Great Britain, 1989.
8 Wilton J. A review of general practice reports: the need for standardisation. *BMJ* 1990;**300**:851-3.
9 Chisholm JW. The 1990 contract: its history and content. *BMJ* 1990;**300**: 853-6.
10 Urquhart AS. Practice annual reports. *J R Coll Gen Pract* 1987;**37**:148.
11 Newlands M, Davies L. The use of performance indicators for immunisation rates in general practice. *Public Health* 1988;**102**:269-73.
12 Jones K, Moon G. *Health, disease and society.* London: Routledge, 1987.
13 Gaze H. FPC defies government over screening targets. *Health Service Journal* 1990;**100**:682.
14 Glasman D. Suspicion takes root in the GP surgery. *Health Service Journal* 1990;**100**:232.
15 Gray J, Jesson D, Jones B. The search for a fairer way of comparing schools' examination results. *Research Papers in Education* 1986;**1**:91-122.
16 Aitken M, Longford N. Statistical modelling in school effectiveness studies. *Journal of Royal Statistical Society A* 1986;**149**:1-43.
17 Goldstein H. *Multi-level models in educational and social research.* London: Charles Griffin, 1987.
18 Bock D, ed. *Multi-level analysis of educational data.* New York: Academic Press, 1989.
19 Jones K, Moon G. A multi-level model approach to immunisation uptake. *Area* 1990;**22**:264-71.
20 Jones K, Moon G, Clegg A. Ecological and individual effects in childhood immunisation uptake. *Soc Sci Med* (in press.)
21 Nelder J, Wedderburn R. Generalised linear models. *Journal of the Royal Statististical Society A* 1972;**135**:370-84.
22 Longford N. *A quasi-likelihood adaption for variance components analysis.* Princeton, New Jersey: Educational Testing Services, 1988.
23 Peckham C. *National immunisation study: factors influencing immunisation uptake in childhood.* London: Institute of Child Health, 1989.
24 Longford N. VARCL: interactive software for variance components analysis. *Professional Statistician* 1986;5:28-32.
25 Morris C. Parameteric empirical Bayes. *Journal of the American Statistical Association* 1983;**78**:47-65.
26 DHSS Performance Indicators Group. *District profiles: a consultation paper.* London: DHSS, 1988.

(Accepted 25 April 1991)

A PAPER THAT CHANGED MY PRACTICE

Problem oriented medical records

A small number of papers changed my practice in managing children with particular conditions, but I can think of only one which did that for every patient whom I saw. I still remember the excitement of reading Lawrence Weed's two part review on "Medical records that guide and teach" in the *New England Journal of Medicine* in 1968. I had increasingly felt dissatisfied with the medical records I had to handle, particularly on patients with complicated problems or long periods of follow up, whose notes became progressively more bulky and less informative. I had also groped towards the idea that identifying the patient's problems should be the basis of medical records. Weed illuminated what was wrong with current records, and persuasively argued for the problem oriented medical record (POMR), where the numbered and dated problem list, prominent at the front of the notes, provides an index and table of contents as well as a concise summary of the patient's history. The problem numbers and titles also provide headings to be used in the progress notes so that everywhere in the record it is clear what is being talked about and the progress of a particular problem can easily be followed in looking through extensive notes. Perhaps the most telling features of Weed's paper were the illustrations comparing the notes of patients before and after a resident had introduced the problem oriented style.

Twenty three years later, Weed's paper still provides an accurate criticism of much medical record keeping today. Some of his ideas—about computerised history taking, for example—seem like yesterday's "Tomorrow's World," but others, like his concern with medical audit, which he rightly believed would be greatly helped by POMR, were far ahead of their time.

For my own practice I quickly adopted the problem oriented medical record and I never abandoned it. It made the notes an actual pleasure to keep, and it changed my approach to children with multiple and complex problems. POMR became a fashionable topic in the 1970s; many hospitals adopted it, but many institutions and doctors who say they use the system do little more than keep traditional style records, with a problem sheet at the front, which is poorly composed or not filled in at all. The full benefits of POMR have not been realised, in more informative records, in patient care, or in medical education.—ROGER ROBINSON, *formerly professor of paediatrics, London; now associate editor, BMJ*

Weed LL. Medical records that guide and teach. *N Engl J Med* 1968;**278**: 593-9,652-8.

THE MEMOIR CLUB

The dinner in our honour was rather an ordeal; there were 12 speeches (all translated) and 12 toasts. The situation was saved by two factors: if anyone spoke for more than two minutes his further remarks were drowned in tumultuous applause which persisted until he sat down; and my neighbour, one of the many lady doctors, most kindly explained the drinking convention to me at the start: "There will be many toasts," she said; "the first glass of vodka must be drained in one motion; thereafter never let your glass become empty again or it will be instantly filled." Twelve sips from the top of the glass can just be managed. We met Harold Wilson at the embassy, over with a friend on business, and on the way home found that they had both bought balalaikas which they were attempting to play and to which they sang any Russian words, however inappropriate, which they could read from the notices around them, revealing a side to the nature of our future Prime Minister which I had not been led to expect.

Our night train to Leningrad on the Red Arrow was a further unusual experience. We travelled with another British medical delegation of three, there to study Soviet thoracic surgery. Jana and their interpreter had brought a great basket filled with bottles of beer and the doctor from the British embassy arrived with his "electrocardiogram" which turned out to be a leather case containing a bottle of whisky and a bottle of vodka. The sleepers on the Red Arrow are wide and we sat on the bunks on either side while the Russians sang gay sounding little songs about such things as children dying in the snow. Passers by along the corridor gathered with the provodnik at the doorway and gradually filtered in, everyone moved up a bit for them to sit down and more people filled the doorway. The songs grew louder, everyone joined in. Presently it was the British turn and song followed song. Those who have not heard Sir Hedley Atkins (later president of the Royal College of Surgeons) rendering "Red Hot Momma" at the top of his voice in a crowded railway carriage in the dead of night somewhere between Moscow and Leningrad have missed one of the greater experiences that life has to offer.