THEORY AND METHODS

Confounding by indication in non-experimental evaluation of vaccine effectiveness: the example of prevention of influenza complications

E Hak, Th J M Verheij, D E Grobbee, K L Nichol, A W Hoes

J Epidemiol Community Health 2002;**56**:951–955

Randomised allocation of vaccine or placebo is the preferred method to assess the effects of the vaccine on clinical outcomes relevant to the individual patient. In the absence of phase 3 trials using clinical end points, notably post-influenza complications, alternative non-experimental designs to evaluate vaccine effects or safety are often used. The application of these designs may, however, lead to invalid estimates of vaccine effectiveness or safety. As patients with poor prognosis are more likely to be immunised, selection for vaccination is confounded by patient factors that are also related to clinical end points. This paper describes several design and analytical methods aimed at limiting or preventing this confounding by indication in non-experimental studies. In short, comparison of study groups with similar prognosis, restriction of the study population, and statistical adjustment for dissimilarities in prognosis are important tools and should be considered. Only if the investigator is able to show that confounding by indication is sufficiently controlled for, results of a non-experimental study may be of use to direct an evidence based vaccine policy.

The health economic impact of influenza epidemics is considerable.[1–3] In most western countries, the use of inactivated influenza vaccines by vulnerable patient groups is advocated to prevent complications.[4] However, uptake of the vaccine remains low, especially in those who need it most.[4–6] Disbelief in the vaccine's effects on clinical outcomes relevant to the individual patient—that is, post-influenza complications—may be one of the major reasons for disappointing immunisation rates.[3 4 6–8]

THE PREVENTION OF INFLUENZA COMPLICATIONS BY VACCINATION: RANDOMISED CONTROLLED TRIALS

The clinical effects of influenza vaccines on reduction of major symptomatic events or death should preferably be studied in phase 3 randomised controlled trials (RCT).[9] Provided that the sample size is large enough, randomised assignment of patients to vaccine or placebo enables valid assessment of vaccine effects through comparing the occurrence of outcomes in both patient groups with similar prognosis. Such trials can be conducted among various segments of the patient population and may give insight into positive as well as negative clinical consequences of immunisation in daily practice. Results of large enough trials in which the primary end point is a clinical outcome rather than a surrogate end point (for example, immune response) provide crucial information on the true impact of these preventive measures and are best suited to guide healthcare decisions.[10 11]

However, scientists face many obstacles when planning a RCT for clinical evaluation of influenza vaccines. Firstly and foremost, as the incidence of influenza related complications or adverse effects is low these trials would entail great expense because large numbers of patients are required.[1 12] Secondly, several influenza seasons may need to be observed as the virulence of circulating influenza viral types is highly variable and unpredictable.[1 6 13] Finally, once the vaccine has been licensed ethical concerns may be raised to further evaluate its effectiveness in placebo controlled studies, especially when persons at high risk for complications are involved. Because of these limitations, post-licensing or phase 4 studies evaluating the vaccine's clinical effectiveness or safety usually use a non-experimental approach, notably a case-control or cohort design.[9] The vaccine's effectiveness is interpreted as the percentage reduction in risk of influenza associated complications attributable to vaccination, given in percentage by 1−RR in cohort studies or 1−OR in case-control studies.[3] The main difference between experimental and non-experimental designs lies in the absence of random allocation of the intervention, for example, vaccination, by the investigator.

EFFECTIVENESS OF INFLUENZA VACCINATION: NON-EXPERIMENTAL STUDIES

One of the important problems encountered in non-experimental evaluation of intended drug effects is the "natural" presence of incomparability of prognosis among subjects receiving the drug and those who do not.[14] In non-experimental influenza vaccine studies, the vaccine group typically comprises patients with more severe disease or (perceived) higher risk, either as a result of self selection or physician preference, than the non-vaccinated (control) group.[15 16] In contrast, those with a contraindication for the intervention will usually be found in the control group only. Thus, selection of exposure is confounded with patient factors, both clinical and non-clinical, which are also related to (detection of) the outcome. This phenomenon may equally apply to qualitative (absence/presence) as well as quantitative (dosing schedule) aspects of exposure and is usually referred to as "confounding by (contra)indication" or "channelling".[14 17 18] Crude, unadjusted, results of non-experiments may therefore lead to invalid inference regarding influenza vaccine effectiveness and potential side effects—that is, underestimation of both beneficial and adverse effects in most circumstances. The obligation of the investigator is to design and analyse the study in such a way that reduction or removal of this type of bias can be achieved.

PREVENTION OF CONFOUNDING BY INDICATION: STUDY DESIGN ISSUES

Preventing or limiting confounding by indication can be achieved in the design and data analytical phase of case-control and cohort studies (see also box 1). In designing

Box 1 Methods to reduce confounding by indication

Design methods
- Comparison of groups with similar prognosis (for example "natural experiment" or use of historical controls)
- Restriction or stratification of study population (for example, age strata, gender, current/inactive disease)
- Individual matching of exposed and non-exposed into main prognostic strata ("quasi-experiment")

Statistical methods
- Statistical control of confounding factors in multivariable regression model
- Subclassification of patients on levels of the propensity score
- Pseudo-randomisation on levels of instrumental variables

a non-experimental study of vaccine effectiveness, valid inferences on preventive effects can be drawn in those situations in which patient groups are compared who have similar indications but have undergone different interventions. These designs could be viewed as "natural experiments". Hypothetically, patients receiving the influenza vaccine because their general practitioner (GP) believes in it and is able to organise the intervention programme (intervention group) could be compared with a group of patients listed with a GP who does not immunise his patients against influenza (control group). Such comparison groups may however be difficult to identify in one healthcare system. Another, less preferred, design option constitutes an ecological study in which vaccine effects among patients residing in different areas are compared. Similarity of ecological comparison groups highly depends on distribution of patient characteristics in different areas. In this respect, a design in which the incidence of influenza associated complications of a historical control group of patients before the introduction of the influenza vaccine is compared with the incidence of such complications in patients after its introduction (intervention group) in one area may be a better option. Such a design, however, risks the incomparability of influenza seasons.

Alternatively, the study domain could be restricted to patients with a more or less similar prognosis such as institutionalised elderly patients.[19] Strict admission criteria could however limit the generalisibility and applicability of results to other segments of the population, while incomparability of comparison groups and residual confounding may persist. Stratification of the study population on levels of important confounding variables, like for example age, and within stratum comparisons also enhances internal validity.[20]

Another option consists of individual pair matching of vaccinated and non-vaccinated subjects within strata of important prognostic variables sometimes referred as "quasi-experiment". This technique was used in a non-experimental evaluation of the effects of placement of ventilation tubes and proved to reduce confounding bias.[21] The design of a quasi-experiment is, however, costly as it requires sufficiently large numbers of patients within each stratum. Except for restriction and stratification, to our knowledge none of the other design options mentioned above has been applied in non-experimental evaluation of currently used influenza vaccines.

Prevention of confounding by indication: data analytical issues

Independent of the study design, statistical adjustment for dissimilarities in prognostic factors between the patient groups receiving and not receiving the vaccine can be applied to improve validity.[1 3 22–24] A prerequisite is that valid and precise data are obtained through the design used to estimate the patient's prognosis without too many missing data. In other words, to optimise statistical adjustment, the prognosis of each patient should be measured by as many valid indicators as possible to permit adjustments afterwards. In primary care, for example, the presence of current disease as indicated by presence of GP consultations in the year preceding the study, also referred as "active patient", is essential to permit valid adjustment of potential confounding. In the ideal situation in which all prognostic patient features can be measured, the exact degree of bias can be quantified and used to draw valid conclusions from the data. In practice, this is usually impossible because of cost restrictions and difficulty, and in that case residual confounding or hidden bias cannot be ruled out. However, although in many non-experimental studies residual confounding may be present, it can be shown that there are limits to the extent of mathematical explanation by this unmeasured confounding. Its putative effects mainly depend on the expected prevalence of the unobserved variable(s), and its associations with vaccination and outcome. Investigators should therefore always reflect on the potential magnitude of the impact of such bias on the effectiveness estimate for example by using sensitivity analysis.[17 26]

In general, three main methods for statistical adjustments can be applied: (1) statistical control of confounding variables in a multivariable regression model[14 18]; (2) subclassifying or matching patients on levels of a so called "propensity score"[17 25–27]; and (3) the use of an instrumental variable to enable statistical pseudo randomisation and to account for any residual confounding.[28]

The first option is commonly used and comprises several steps: identification of confounders in the dataset, univariate stratification of exposure groups on levels of the confounder to estimate the vaccine effectiveness estimate adjusted for this single variable (for example, age) and multivariable control including confounding variables that collectively influence the estimated relation between exposure and outcome in the modelling procedure.

A method to optimise statistical adjustment for confounding by indication in non-experimental studies, notably when the number of prognostic variables is large, has been proposed by Rubin and Rosenbaum. They introduced the "propensity score" method.[17 25–27] This score is the conditional probability of exposure to a treatment given a set of observed variables that may influence the decision to vaccinate. The propensity score can be derived from a multivariable logistic regression analysis in which those variables that are statistically significant associated with exposure (for example, vaccination) are included. Obviously, the outcome variable should not be included as a covariate. A higher score indicates a higher probability of receiving the vaccine. Subclassification of subjects on levels of this single variable or including this variable as a single covariate in a multivariable regression model tends to balance all of the observed variables, but not the unobserved.[17 25–27] The use of this score and matched sampling will also implicitly incorporate any interactions among confounders. Thus, this technique enables the investigator to assess the association of vaccination with specific outcomes in patients with a more or less equal probability of receiving the vaccine. Discriminant matching for multivariate normal covariates as described by Cochran[29] and the use of a "confounder score" as proposed by Miettinen are related techniques.[30]

To overcome the potential lack of balance on unobserved prognostic indicators (for example, health behaviour), the instrumental variable method has been suggested. This technique originates from the field of econometrics and has so far not been extensively used in medical research. In short, patients are subdivided according to levels of a covariate that is associated with the exposure, but not associated with the

Table 1 Crude and adjusted odds ratios for an acute episode of low respiratory tract or cardiac disease or death during an influenza epidemic in vaccinees and non-vaccinees

Study population and analysis	Adjusted for:	Odds ratio (95% CI)
Adult patients (18–102 y, n=1696) Conventional control: MLR*	Crude value	1.14 (0.84 to 1.55)
	+ age (in years)	0.87 (0.64 to 1.20)
	+ disease (asthma/COPD)	0.82 (0.59 to 1.13)
	+ GP visits (in number)	0.76 (0.54 to 1.05)
	+ remaining factors	0.76 (0.54 to 1.06)
Elderly patients (65–102 y, n=630) Conventional control: MLR	Crude value	0.57 (0.35 to 0.93)
	+ age (in years)	0.56 (0.35 to 0.92)
	+ disease (asthma/COPD)	0.53 (0.32 to 0.87)
	+ GP visits (in number)	0.50 (0.30 to 0.83)
	+ remaining factors	0.50 (0.29 to 0.83)
Younger patients (18–64 y, n=1066) Conventional control: MLR	Crude value	1.27 (0.84 to 1.94)
	+ age (in years)	1.11 (0.73 to 1.70)
	+ disease (asthma/COPD)	1.08 (0.70 to 1.66)
	+ GP visits (in number)	0.94 (0.61 to 1.47)
	+ remaining factors	0.94 (0.60 to 1.45)
Quasi-experiment (18–64 y, n=676) Conventional control: MCLR†	Matched crude value	0.90 (0.53 to 1.52)
	+ age/ disease/GP visits + remaining factors	0.89 (0.52 to 1.54)
Younger patients (18–64 y, n=1066) Propensity score + MCLR†	Matched crude value	0.87 (0.56 to 1.35)
	+ age/ disease/GP visits + remaining factors	0.86 (0.55 to 1.35)

*MLR, multivariable logistic regression analysis; †MCLR, multivariable conditional logistic regression.

outcome. This pseudo-randomisation may lead to equal distribution of health characteristics in both non-exposed and exposed people and thus prevent potential confounding. For example, McClellan *et al* calculated the distance to the hospital on the basis of zip codes and divided patients into those living within a small area around the hospital and those outside that area in a study on cardiovascular procedures.[28] Distance to the hospital did fulfill the criteria for instrumental variables. Heart catheterisation was more prevalent in the inner circle than the outer circle, and mortality rates were similar. This was in contrast with their finding using conventional control for confounding in which mortality rates seemed higher in patients who underwent the surgical procedure. As the validity of this second method should be evaluated in other medical studies and instrumental variables may be hard to identify, we will not further elaborate on this statistical procedure.

The presence of confounding by indication in non-experimental evaluation of influenza vaccination and some of the above mentioned tools to reduce its impact are discussed in more detail on the basis of data derived from a recent study by our group.

AN EXAMPLE: INFLUENZA VACCINE EFFECTIVENESS IN ADULT PATIENTS WITH PULMONARY DISEASE

We examined the effect of influenza vaccine on the incidence of influenza associated complications in 1696 adult patients with chronic obstructive pulmonary disease (COPD) or asthma during the 1995/96 influenza A epidemic.[31] The study was a one season prospective cohort study using the medical database of the Utrecht General Practitioners Network. GP patient records were reviewed for all study subjects. As a first design approach to limit confounding by indication, vaccinated and non-vaccinated patients with pulmonary disease were compared rather than vaccinated patients and controls from the community. The study population was restricted to those with an indication for vaccination according to the guidelines of the Dutch Health Council. In table 1 we give crude and adjusted effectiveness estimates using the conventional control of confounding by multivariable logistic regression analysis. Despite restriction of the study population, crude results seem to suggest that the vaccine is ineffective and may even lead to complications (odds ratio (OR) 1.14). If the effectiveness estimate of 50% reduction that we observed among only 630 elderly patients is hypothesised to be present in the total study group, the size (n=1696) gives a power over 80% to detect a reduction from 10% to 5% of complications with a vaccination rate of 66%. So this cannot be explained by lack of statistical power only. However, further statistical adjustments notably for age, disease, and GP visits resulted in striking changes of the effectiveness estimate to a relative risk of 0.76 suggesting an overall vaccine effectiveness of 24% in this population—a relative parameter change of 33%. Confounding therefore might explain the observed associations. Addition of other covariates, including the GP, in the final model did not substantially change the vaccine effectiveness estimate.

Most probably the adjustments were still incomplete. More precise measurements of disease severity such as pulmonary function, atopy, or hyper-reactivity were not available. Therefore, a second approach to limit confounding consisted of subdividing the whole study population into two age strata (⩾65 years, 18–64 years) in which prognosis of vaccinees and non-vaccinees within each age stratum is less deviant (see also table 1). Apart from issues of modification of the effects of the vaccine by age, which is beyond the scope of this article, with this approach, statistical adjustments for the same confounding factors resulted in smaller relative parameter changes of 12% and 26%, respectively, in both age categories.

Key points

- In non-experimental effectiveness studies, selection for drugs, for example, vaccination, can be confounded by patient factors that are also related to clinical end points.
- Comparison of similar groups, restriction, or matching are design methods to reduce confounding by indication.
- Regression analysis, subclassification on propensity scores and pseudo-randomisation with instrumental variables are statistical methods to reduce confounding by indication.

This suggests that stratification or age restriction may further reduce residual confounding. Still, inferences on the two age subgroups should be made with caution. In the elderly population, a substantial and statistically significant reduction in the outcome rate was observed even without controlling for confounding (OR 0.57, 95% confidence intervals (CI) 0.35 to 0.93). Addition of prognostic factors into the multivariate model led to a further increase in the estimate of vaccine effectiveness indicating some residual confounding after stratification. However, in the working age adults the crude odds ratio was well above 1.0 and despite adjustment for the available prognostic indicators we could not demonstrate a significant reduction (OR 0.94, 95% CI 0.61 to 1.47). This suggests that results of restricted populations are not necessarily applicable to other segments, in this case younger patients. Also, in this segment the GP did not independently confound the association. Although vaccination rates of the seven group practices were somewhat different (range 56% to 68%, p=0.08), complication rates were almost similar (range 8% to 13%, p=0.24). Because Neuzil and colleagues showed considerable impact of influenza in a younger group of women[6] and we have shown that in the Netherlands the current influenza target group comprises at least 40% of high risk persons under 65 years of age,[32] we further examined potential confounding in this particular age group.

As a third approach to limit potential confounding by indication in the original design, we used the data of this younger age group (18–64 years) in a "quasi-experiment". Firstly, we identified the three main prognostic factors: age (five year age category), underlying pulmonary disease (asthma or COPD) and GP visiting rate (0, 1–2, and ⩾3 visits). Next, we classified each subject, vaccinated or non-vaccinated, into one of the 54 combinations of these factors. Within each stratum we then randomly sampled from either the vaccinated or the non-vaccinated group as many patients as were available in the comparison group with the lowest number of subjects. For example, if five vaccinated and two non-vaccinated patients were between 20 and 24 years old, had asthma and consulted the GP five times in the preceding year, we sampled two patients at random from the exposed group to form a stratum matched group. In all, 390 patients (37%) were excluded from the original study population (n=1066) and 676 patients were available for the quasi-experiment. After this matching procedure it seemed that the vaccine reduced the occurrence of outcomes by 11%, after adjustments for the main confounders and remaining covariates (that is, health insurance, gender), but the estimate was not statistically significant (see table 1). Only minor changes were observed after statistical adjustment, suggesting that confounding by differences in the known prognostic factors was largely removed. A major limitation may prohibit the use of the above mentioned "quasi experiment". Pair matching is time consuming and can considerably reduce the power of the study as numbers of matched patients in separate strata become small. In our example 37% of the initial study population had to be excluded. To avoid these issues, we finally applied analytical control of confounding by using the "propensity score".

In our example, we used the 1066 patients aged between 18 and 64 years to calculate the probability score of being vaccinated. Our final multivariable logistic regression model with the dependent variable vaccination included age, underlying disease, number of GP visits, gender, and health insurance. We then categorised the propensity score into quintiles and matched vaccinees and non-vaccinees on levels of the probability to be vaccinated. In the multivariable conditional logistic regression analysis we matched on the categorised levels of the score and calculated crude and adjusted odds ratios of vaccination for the outcome. The overall adjusted odds ratio of 0.86 seems to suggest a 14% reduction of complications resulting from the vaccine. The finding of the "quasi-experiment" in which stratum matched pairs of vaccinees and non-vaccinees were compared was validated by this statistical method. As was expected, 95% confidence intervals were smaller, but point estimates were nearly the same. The latter techniques changed the effectiveness estimate from a crude estimate of −27% in the original design to 11% and 14% using the "quasi-experiment" and "propensity score", respectively; relative parameter changes of more than 30%. In addition, the propensity score method resulted in slightly smaller 95% confidence intervals than the conventional adjustment. Although our study lacked adequate power to demonstrate a statistically significant reduction of outcomes resulting from the vaccine, the adjusted effectiveness point estimates are compatible with a statistically significant 11% reduction of outpatient visits for respiratory disease in elderly lung patients as observed by Nichol and colleagues.[33]

CONCLUSION

Randomised allocation of vaccine or placebo is the preferred method to assess the effects of the vaccine on clinical outcomes relevant to the individual patient. In the absence of phase 3 trials using clinical end points, alternative non-experimental designs to evaluate vaccine effects or safety are often used. The application of these designs may, however, lead to invalid estimates of vaccine effectiveness or safety. As patients with poor prognosis are more likely to be immunised, selection for vaccination is confounded by patient factors that are also related to clinical end points. This paper describes several design and analytical methods aimed at limiting or preventing this confounding by indication in non-experimental studies. In short, comparison of study groups with similar prognosis, restriction of the study population, and statistical adjustment for dissimilarities in prognosis are important tools and should be considered. Only if the investigator is able to show that confounding by indication is sufficiently controlled for, results of a non-experimental study may be of use to direct an evidence based vaccine policy.

Authors' affiliations
E Hak, Th J M Verheij, D E Grobbee, A W Hoes, Julius Centre for General Practice and Patient Oriented Research, University Medical Centre Utrecht, Netherlands
K L Nichol, Medicine Service, Veterans Affairs Medical Center and University of Minnesota, Minneapolis, USA

Funding: UMC Utrecht and The Netherlands Asthma Foundation (no 97.51).

Conflicts of interest: none

Correspondence to: Dr E Hak, University Medical Center Utrecht, Julius Center for General Practice and Patient Oriented Research, Location Stratenum, PO Box 85060, 3508 AB Utrecht, Netherlands; E.Hak@med.uu.nl

Accepted for publication 18 January 2002

REFERENCES

1 **Nichol KL**, Margolis KL, Wuorenma J, *et al*. The efficacy and cost-effectiveness of vaccination against influenza among elderly persons living in the community. *N Engl J Med* 1994;**331**:778–84.
2 **Barker WH**, Borisute H, Cox C. A study on the impact of influenza on the functional status of frail older people. *Arch Intern Med* 1998;**158**:645–50.
3 **Nichol KL**, Wuorenma J, Von Sternberg T. Benefits of influenza vaccination for low-, intermediate, and high-risk senior citizens. *Arch Intern Med* 1998;**158**:1769–76
4 **Fedson DS**, Hirota Y, Shin H, *et al*. Influenza vaccination in 22 developed countries: an update to 1995. *Vaccine* 1997;**15**:1506–11.
5 **Watkins J**. Effectiveness of influenza vaccination policy at targeting patients at high risk of complications during winter 1994–5: cross sectional survey. *BMJ* 1997;**315**:1069–70.
6 **Neuzil KM**, Reed GW, Mitchels EF, *et al*. Influenza-associated morbidity and mortality in young and middle-aged women. *JAMA* 1999;**281**:901–7.
7 **Essen GA van**, Kuyvenhoven MM, Melker RA de. Compliance with influenza vaccination. Its relation with epidemiologic and sociopsychological factors. *Arch Fam Med* 1997;**6**:157–62.
8 **Frank JW**, Henderson M, McMurray L. Influenza vaccination in the elderly, I: determinants of acceptance. *Can Med Assoc J* 1995;**132**:371–5.
9 **Clemens J**, Brenner R, Rao M, *et al*. Evaluating new vaccines for developing countries. *JAMA* 1996;**275**:390–97.
10 **Haddix AC**, Teutsch SM, Shaffer PA, *et al*, eds. *Prevention effectiveness. A guide to decision analysis and economic evaluation*. New York: Oxford University Press, 1996.
11 **Fleming TR**, DeMets DL. Surrogate end points in clinical trials: are we being misled? *Ann Intern Med* 1996;**125**:605–13.
12 **Rothbart PH**, Kempen BM, Sprenger MJW. Sense and nonsense of influenza vaccination in asthma and chronic obstructive pulmonary disease. *Am J Respir Crit Care* 1995;**151**:1682–6.
13 **Glezen WP**. Serious morbidity and mortality associated with influenza epidemics. *Epidemiol Rev* 1982;**4**:25–44.
14 **Grobbee DE**, Hoes AW. Confounding and indication for treatment in evaluation of drug treatment for hypertension. *BMJ* 1997;**315**:1151–4.
15 **Rodriques LC**, Smith PG. Use of the case-control approach in vaccine evaluation: efficacy and adverse effects. *Epidemiol Rev* 1999;**21**:56–72.
16 **Kramarz P**, DeStefano F, Gargiullo PM, *et al*. Influenza vaccination in children with asthma in Health Maintenance Organizations. *Vaccine* 2000;**18**:2288–94.
17 **Rubin DB**. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med* 1997;**127**:757–63.
18 **Greenland S**, Neutra R. Control of confounding in the assessment of medical technology. *Int J Epidemiol* 1980;**9**:361–7.
19 **Gross PA**, Hermogenes AW, Sacks HS, *et al*. The efficacy of influenza vaccine in elderly persons. A meta-analysis and review of the literature. *Ann Intern Med* 1995;**123**:518–27.
20 **Fedson DS**, Wajda A, Nicol JP, *et al*. Clinical effectiveness of influenza vaccination in Manitoba. *JAMA* 1993;**270**:1956–61.
21 **Schilder AGM**, Hak E, Straatman H, *et al*. Long-term effects of ventilation tubes for persistent OME in children. *Clin Otolaryngol* 1997;**22**:423–9.
22 **Nichol KL**, Baken L, Nelson A. Relation between influenza vaccination and outpatient visits, hospitalization, and mortality in elderly persons with chronic lung disease. *Ann Intern Med* 1999;**130**:397–403.
23 **Ahmed AH**, Nicholson KG, Nguyen-van-Tam JS. Reduction in mortality associated with influenza vaccine during 1989–90 epidemic. *Lancet* 1995;**346**:591–5.
24 **Ohmit SE**, Monto AS. Influenza vaccine effectiveness in preventing hospitalization among the elderly during influenza type A end type B seasons. *Int J Epidemiol* 1995;**24**:1240–8.
25 **Rosenbaum PR**, Rubin DB. Reducing bias in observational studies using subclassification on the propensity score. *J Am Stat Assoc* 1984;**79**:516–24.
26 **Connors AF**, Speroff T, Dawson NV, *et al*. The effectiveness of right heart catherization in the initial care of critically ill patients. *JAMA* 1996;**276**:889–97.
27 **Joffe MM**, Rosenbaum PR. Invited commentary: Propensity score. *Am J Epidemiol* 1999;**150**:327–33.
28 **McClellan M**, McNeil JP. Newhouse: Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA* 1994;**272**:859–66.
29 **Cochran WG**, Rubin D. Controlling bias in observational studies: a review. *Sankya* 1973;**35**:417–46.
30 **Miettinen OS**. Stratification by a multivariate confounder score. *Am J Epidemiol* 1976;**104**:609–20.
31 **Hak E**, Essen GA van, Buskens E, *et al*. Is immunising all patients with chronic lung disease in the community against influenza cost effective? Evidence from a general practice based clinical prospective cohort study in Utrecht, the Netherlands. *J Epidemiol Community Health* 1998;**52**:120–5.
32 **Hak E**, Essen GA van, Stalman WAB, *et al*. Improving influenza vaccination coverage among high-risk patients: a role for computer-supported prevention strategy? *Fam Pract* 1998;**15**:138–43.
33 **Nichol KL**, Baken L, Nelson A. Relation between influenza vaccination and outpatient visits, hospitalization, and mortality in elderly persons with chronic lung disease. *Ann Intern Med* 1999;**130**:397–403.