

How Real is Intention-To-Treat (ITT) Analysis in Non-Interventional Post Authorization Safety Studies? We Can Do Better

Victor A Kiri^{*,1,2} and Gilbert MacKenzie²

1 Peri-Approval, Clinical Research Services, PAREXEL International, Uxbridge, UK

2 Centre for Biostatistics, University of Limerick, Ireland

*Address correspondence to this author at the Department of Peri-Approval Clinical Excellence, Clinical Research Services, PAREXEL International, Uxbridge, Middlesex UB8 1LZ, UK; Tel: +44 (0) 1895 614760; Fax: +44 (0) 1895 614277;
E-mail: Victor.Kiri@PAREXEL.com

ABSTRACT

Although cohort studies which are based on intention-to-treat (ITT) approach offer a simple design with data which are simpler to analyse and results easier to interpret, such studies also intrinsically assume that any time-varying treatment effect that exists can be adequately estimated by a fixed-effect component. However, such an assumption may not reflect real-life drug use. Reflection of real-life clinical practice is a major strength of epidemiologic safety studies. The failure to properly reflect reality may result in effect under-estimation leading to false and irreproducible conclusions due to exposure misclassification. In effect, the use of nested case-control design is a concession that ITT in cohort design may not be adequate. But the nested design also has its own sources of bias, including confounding by indication. We present an overview of the counter-matched version of the nested case-control, case-crossover, case-in-time, case series and case-cohort designs as alternatives in prospective post-authorization safety studies.

Key words: Cohort study, selection bias, Case-Control design, counter-matching

INTRODUCTION

Pharmacoepidemiological studies are becoming more common in the EU regulatory framework for drug safety assessments, thanks largely to Article 1(15) of Directive 2001/83/EC which defines a post-authorisation safety study (PASS) as either “pharmacoepidemiological study or a clinical trial carried out in accordance with the terms of marketing authorisation, conducted with the aim of identifying or quantifying a safety hazard relating to an authorised medicinal product”. The main reason for their popularity may lie in their non-experimental nature- being essentially observational in practice and hence, reflecting real-life medical practice. However, observational studies are not easy to conduct in the sense of meeting their stated objectives. The main difficulty is largely due to the inherent risk of bias associated with these studies. Hence, appropriate design, efficient analytical strategy and appropriate interpretation of results are vital.

PASS are based almost entirely on the cohort design which entails recruiting all eligible patients with or without exposure to the therapy of interest and subsequently following them for a specified period of time. At the analytic stage, we compare exposed patients with the unexposed. However, by design, the exposure status on each patient is usually determined at baseline or at a specified time point from recruitment, considered as the formal start of follow-up. Thus, an intention-to-treat (ITT) assumption is made in the analysis of the resulting data. ITT analysis is one that is based on the initial treatment intent and not necessarily on the treatment eventually administered. Cohort studies which are based on intention-to-treat approach involve simple designs which offer datasets that are simpler to analyse and its results are easier to interpret. Of course, the ITT assumption suggests any time-varying treatment effect can be adequately estimated by a fixed average. In effect, we assume that exposure occurs continuously and the duration of exposure is identical for all exposed patients. However intermittent exposure is more common in real-life medical practice than may be assumed and the problems associated with simplistic assumptions on exposure classification in pharmacoepidemiological studies have been highlighted elsewhere [1-2].

In a clinical trial, the aim of the ITT assumption is to estimate the effect of the decision to apply the treatment and the analysis considers treatment groups as allocated regardless of compliance. Thus, the analysis provides information about the potential effects of treatment policy rather than on the potential effects of the specific

treatment. By contrast, because treatment allocation is informative in an observational study, adoption of the ITT assumption will mean estimation of the effect of the decision to make the treatment available. Thus ITT assumption on exposure classification may be suitable in a clinical trial PASS where random assignment of therapy can ensure its validity. Indeed, the assumption is effected in clinical trials to avoid the effects of crossover and drop-out which may break the randomization to the treatment groups. However, for randomized safety trials there are potential dangers- poor adherence to treatment will tend to yield inaccurate estimates of the safety of the treatment under study due to exposure misclassification. In observational studies where the absence of randomization makes the potential for confounding our main concern, identification of the actual exposure of the individuals can help to reduce misclassification. Indeed, reflection of actual drug use is perhaps, the most compelling strength of pharmacoepidemiological safety studies. We know that failure to properly reflect reality may mean exposure misclassification which for a safety study may result in effect under-estimation, false and irreproducible conclusions. In effect, the use of the nested case-control design in certain observational cohort studies can be viewed as a form of concession that ITT in observational studies may not be adequate. Furthermore, the need to ensure comparability between the exposed and unexposed patients in an observational cohort design often results in the exclusion of a large number of the study population at the analytical stage- making data collection a costly handicap of the prospective version of the design.

Unfortunately, although modelling treatment as time-dependent can be an option to ITT in observational studies, the approach has its own assumptions, some of which may be difficult to satisfy in a PASS setting where inappropriate application can lead to a more profound form of bias- false association. The validity of the time-dependent treatment model depends on satisfaction of the relatively strong assumption that the reason for change of exposure status is unrelated to the subsequent probability of an event. Thus the channelling of a drug may increase the likelihood of detecting a false adverse effect as more patients receive the drug when they have a poorer prognosis. In any case, time-dependent treatment effects are not that simple to interpret, especially when the issue is about the safety or otherwise of a drug and quite rightly, simple but clear message remains a general preference.

We present an overview of the nested case-control, case-crossover, case-in-time, case series and case-cohort designs which we propose as alternatives in

pharmacoepidemiological PASS involving prospectively collected data. In particular, we (1) highlight the potential of counter-matching for improving the efficiency of the nested case-control design and (2) describe how the case-cohort design can be adopted to extend safety assessment activities beyond the unblinding phase of post approval clinical trials.

THE ALTERNATIVES

Counter-matched nested case-control design:

The nested case-control design offers a simple method for avoiding unreasonable assumptions in the evaluation of time-dependent treatment effect [3]. For a given event of interest, the design involves identifying suitable controls (i.e. patients free of the event) at each time of an event which are then matched to the case (i.e. patient with the event) based on a prior specified confounding factors. The process of control selection is usually by random sampling and controls can be future cases. The results of the nested case-control design are easy to interpret [4-6]. Although the main strength of the design rests largely on the appropriateness of the controls that are matched to the cases, not much effort is generally made to evaluate its efficiency.

Matching controls to cases on a confounding variable can improve the precision of the comparison of exposure groups although the effect of that variable cannot then be estimated. Whilst we do not match on the exposure variable, we still cannot estimate exposure effect, if in the matched data, all the cases are of the same exposure status and same will be true if all the controls are of the same exposure status or the status is the same on each matched pair. In other words, effect estimation is based entirely on the off-diagonal data of the resulting stratum-based 2x2 tables in the conditional logistic regression. Consequently in theory, the more such matched pairs our sampling can generate the more improvement we can expect on efficiency.

As illustration, suppose for any particular risk set have N_e and N_u as number of exposed and unexposed subjects respectively from which we are to draw m controls. Then the nested case-control set will contain $n=m+1$ subjects and suppose the split of these n subjects between exposed and unexposed is n_e and n_u .

When controls are drawn by simple random sampling of the risk sets, the resulting data may be comprised of an uneven split of exposed and unexposed subjects which can affect the efficiency of the design. To optimise efficiency, statistical theory suggests counter matching cases and controls on exposure status based on information in the full

cohort [7]. Where this approach is adopted, data collection can be limited to only the cases and the sample of patients matched to the cases on a set of baseline characteristics that are specified a priori. According to statistical theory, analysis of this small cohort is capable of reproducing the results from a full cohort analysis free of time-varying exposure effect and confounding variables.

When sampling in this way the contribution of each risk set to the partial log likelihood must be adjusted to reflect the fact that the exposure in the sample is different from the exposure distribution in the risk set. The modified log partial likelihood contribution takes the form:

$$\log \left(\frac{(W\theta)_{\text{for case}}}{\sum_{\text{case-control set}} (W\theta)} \right)$$

Where W are risk weights for each subject which compensate for the sampling.

These weights take the values: N_e/n_e for an exposed subject and N_u/n_u for an unexposed. The choice of weights depends only on exposure status and not upon whether the subject is a case or a control.

Thus, in the prospective framework, it will not be necessary to collect all data on all patients beyond the baseline, which should translate to cost savings. Indeed, the counter-matched nested case control design can be particularly attractive in a prospective study where (1) the interest is on either a particular event or a prior specified collection of events, (2) we do not wish to collect a lot of data and/or (3) the cost of a normal registry/cohort will be too high.

Nevertheless, the nested case-control design is generally inadequate if the reason for the treatment may be associated with the particular outcome of interest, in which case the design may present problems such as confounding by indication [8], susceptibility bias [9] and/or channelling bias [10]. Since matching controls to cases on a confounding variable can improve the precision of the comparison of exposure groups, we can reduce the bias due to confounding by incorporating a balancing statistic as one of the matching variables. Such a balancing tool can be derived using either propensity scores [11-14] or the Heckman selection models [15-18].

Case-crossover Design:

In situations where it may be difficult to find adequate comparable controls to the cases in a nested case control design, a case-crossover design may be preferable.

Cross-over is a term used to describe experiments (i.e. clinical trials) in which all subjects serve in the treatment and placebo groups at different times of the study. The case-crossover design considers the cases themselves as the best representatives of the population base that produces them [19]. Thus only cases with dynamic exposure status (i.e. periods of exposure and non-exposure) are included in the study, with each serving as his/her own control. Indeed, the controls are the same cases but at earlier times and as such, the design is based on subject-matched sampling. Exposure frequencies from the case window (i.e. current period) and the control window (i.e. reference period) are compared through a matched odds ratio.

To illustrate how easy it is to apply, we consider the following as the number of patients such that the observed odds of exposure is: a_i/b_i divided by N_{ei}/N_{ui} :

	Exposed	Unexposed	Total
Case	a_i	b_i	$a_i + b_i$
Control	N_{ei}	N_{ui}	$N_{ei} + N_{ui}$
Total	$a_i + N_{ei}$	$b_i + N_{ui}$	T_i

Then we can obtain the Mantel and Haenszel Odds ratio as: $OR_{MH} = \frac{\sum_i a_i N_{ui} / T_i}{\sum_i b_i N_{ei} / T_i}$

This odds ratio is a measure of association between exposure and risk which can be computed for several subgroups and it is generally considered the best statistical adjustment for confounding variables with time-constant effects. Indeed, the design is particularly good for reducing the effect of unmeasured confounding [20].

However, an inherent assumption in the case-crossover design is a constant subject-specific exposure distribution: exposure should have the same distribution in case and control times [21]. Thus, bias may result when the assumption is not valid. Although a bidirectional sampling of control windows could eliminate part of the bias from exposure time trend [22], the choice of a control window may still be troublesome because of problems of induction periods and carry-over effects.

The case-crossover design is useful for assessing the transient effect of risk factors for diseases with acute onset. However it is not appropriate for chronic diseases or diseases with carry-over risk effects [23]. The design offers the following advantages:

- Control selection bias is eliminated as the cases act as their own controls. Thus the design can be very useful for studies where choice of appropriate controls is

a problem or where ethical issues prevent the involvement of control patients. It is particularly useful for common vaccine studies.

- In the prospective framework, we can restrict collection of detailed information to the cases with dynamic exposure status.
- Fewer resources are needed since there is no need to collect information on time-constant factors or a separate group of controls.
- Bias due to confounding by indication will be eliminated or reduced if the reasons for prescribing are relatively constant over the period of the study.

However, the design also has the following disadvantages:

- Due care is required in the interpretation of the risk ratio estimates.
- Self matching can decrease precision if the number of concordant pairs is high (i.e. the exposure is long lasting and bridges periods).
- There is always the danger that due to variations in disease severity, the cases may form a genuinely distinct group of patients from the population they purport to represent.
- The odds ratios estimates, being based solely on the cases may also represent the natural increase in drug use over the time and not just the increase associated with the event. This is because by design, a portion of the effect of any natural time trend in medication use will be included in the estimates.
- It is not suitable if exposure effect is not transient

Case-in-time Design:

This design is an extension of the case-crossover design as it takes into account, changes in drug utilization over time, thereby separating the effects of the drug and time [24-25]. The effect of time can result from changes in disease severity or the spectrum of stages of a disease which can be an important confounding factor. The design incorporates both the case-crossover and nested case-control designs. From the case-crossover design, we select all cases with dynamic exposure status but from the nested case-control design, we only select all the matched control subjects with dynamic exposure status which we call the “controls”. The resulting patients are then self-matched based on the current and reference periods as in the case-crossover design to constitute the pairs.

The case-time-control design is based on two main assumptions:

- That the odds ratio from the case-crossover design is the product of the odds ratio due to exposure and the odds ratio due to time trend in exposure prevalence
- That the time trend is the same among the cases and “controls”

For our estimations, we assume conditional independence of exposure within each pair. This allows us to use the conditional logistic regression model to estimate the odds ratios for period and outcome effects. Conditional analysis automatically eliminates the nuisance parameters induced by the subject effects which inherently includes the effects of disease severity. With exposure as an independent variable and period as the dependant variable we estimate the portion of the risk associated with the period effect from the “controls” whilst from the cases, we obtain the estimate associated with both period and exposure. Thus, to obtain the odds ratio estimate for drug effect we divide the odds ratio estimate from the cases by the odds ratio estimate from the “controls”.

The case-time-control design offers the following advantages:

- It permits separation of the effect associated with the drug from that of disease severity, even if this severity is not measured
- Information bias may turn out to be of a less concern than in the case-crossover design: indeed if exposure measures are obtained the same way for cases and controls, any bias induced by time should cancel out
- It is particularly suitable for assessing transient effect of treatment where the outcome is an acute onset disease. However, it can be used to assess exposure effect over a window of up to a year.
- It is best when the exposure trends can be estimated with great precision and the degree to which it provides unbiased results is dependent upon how well these trends are estimated
- In the prospective framework, we can restrict collection of detailed information to the cases and matched controls with dynamic exposure status.

However, the design also has the following disadvantages:

- Selection bias is possible where exposure change among cases is different from exposure change among the “controls”
- If severity, which is associated with drug use also increases (or decreases) within subjects over time, and does so differently for case and control subjects, what is

believed to be the residual effect associated with the drug could well remain confounded to some extent [21, 24-26]

- It usually produces effect estimates that are systematically lower than those from the case-control design [27]
- The within subjects correlation induced by this design may result in less precise effect estimate. But the apparent loss is an artefact since the case-control approach produces biased estimates of drug effects, confounded by disease severity. The precisions are thus not comparable [28]

Case series Design:

Like that of the case-crossover, this design is also based on cases only but it is particularly suitable for prospective observation where we can restrict collection of detailed information to the cases. Furthermore, the design specifically addresses the problem of intermittent exposures and acknowledges the role of concomitant factors which can influence the hazard rates differentially during periods of exposure and non-exposure [29]. It does this by assuming piecewise constant hazard functions over fixed periods of observation. The null hypothesis is usually that the event rates remain constant from day to day and are not affected by the exposure. The periods of observation are usually specified a priori, beginning with the exposure period which starts from the date of exposure. All the other observation periods are then collectively considered as the baseline periods (i.e. without exposure). For a patient with multiple exposures, it is normal to consider each exposure as constituting a separate set of exposure and baseline periods, provided the patient can be reasonably considered as being at risk in each of the periods. By treating each period as a stratum, hazard ratios are then calculated for events that occur within each stratum of the period of exposure as compared with the baseline periods. The design can offer simple and objectively measurable criteria for the assessment of drug safety [30-32].

Case-cohort Design:

For situations where the interest is on multiple events, we recommend the use of the case-cohort design. The design can facilitate effective assessment of multiple outcomes without much of the problem of correlated effects [33-35]. It is most useful for analysing time to failure in a large cohort in which failure is rare, such as in the study of serious adverse drug effects. The case-cohort design involves randomly selecting a

representative sample of the patients in the full cohort, irrespective of case status (known as the sub-cohort). Relevant baseline information is collected on every member of the sub-cohort and also on those patients in the remainder of the full cohort who experience the events of interest. Another of its distinctive features is that where as each member of the sub-cohort can contribute to the risk set up to the time of being a case or censored, a case outside the sub-cohort will only contribute at the exact time of becoming a case.

In the prospective framework, the case-cohort design can be used to address the problem of intermittent exposure more easily than in a full cohort. We can afford to update the relevant sets of information on members of the corresponding risk set at each event point (i.e. the case and remaining sub-cohort members) which for time-varying exposure, we can treat as a categorical variable in the same way as in the nested case-control design. We see a role for this design for continued assessment of serious adverse drug effects over a longer period of time at the end of on-going Phase IV trials. The sub-cohort can be easily obtained as a random (possibly stratified) sample of the new study population (i.e. the consenting trial population around the time of unblinding) to generate useful prospective data.

To encourage interest in the design as a promising approach to risk estimation, we present the essential elements of its formulation.

Computational Formulation:

In a cohort design, we usually adopt the proportional hazards model to obtain the risk estimates. In doing so, we rely on three main assumptions on the resulting partial likelihood- (1) independent failure times and a censoring process that is independent of that of failure, (2) with an unspecified underlying hazard, no additional information about the risk can be obtained from time intervals with no failure and (3) information contributed by a single risk set does not generally increase by virtue of its size [36-37]. These assumptions suggest considerable economy may be achieved at little cost in the precision of the estimate by comparing “failure” with only a random sample of those at risk at the given time. Insight into the loss of precision to be expected due to the discarded numbers at risk is given by the standard formula for variance of the difference in the means of independent samples.

The same assumptions facilitate consideration of the case-cohort design as an alternative, where the risk set is constituted only by members of a sub cohort. The

results are a pseudo-likelihood which mimics a form of the partial likelihood for the risk estimation and a variance estimation that requires computation of the covariance among the score components arising from the sampling design [38]. Although the score components are likely to be correlated, the design allows for direct estimation of the risk ratio (i.e. ratio of incidence proportions between exposed and unexposed), without the rare-disease assumption associated with case-control [39]. Indeed, a jack-knife approach which is based on estimates of the influence function (i.e. the influence of each individual on the risk and on the overall score) and is computationally more efficient and generalizable to other sampling scheme is now available [34-35]. The new method is concerned with the contribution of the score function of each individual summed over failure times, rather than contributions at each failure time summed over individuals. The variance estimate is identical to a robust variance estimator- robust to misspecification of the model and measures of influence of individual observations can be easily obtained for sensitivity analysis [40-41]. Furthermore, the jack-knife approach also provides an adequate correction for the correlation of the risk estimates which may result from the assessment of multiple outcomes. The statistical software- SAS (version 8.2; SAS Institute; Cary, NC) has within its PHREG procedure, the relevant option for application of the case-cohort design.

CONCLUSION

In pharmacoepidemiological studies, the popular intent-to-treat assumption on exposure classification may be unsuitable. Exposure is seldom continuous in real-life and intermittent drug intake is common practice. For safety studies, such bias can result in effect underestimation and possibly, wrong conclusions and contradictory irreproducible estimates. This is a problem we ignore when we design a simple registry to assess the safety of a given drug. This is in addition to the problem of selection bias which is common in observational studies due to the absence of random assignment of treatment. To reduce the effect of selection bias, it is usual practice to select from the study population, only exposed and unexposed patients that are comparable on key confounding factors. Consequently, we often end up discarding much of the data collected in the registry.

The nested case-control design which is based on a set of matched cases and their controls (i.e. matching on certain key factors) from the registry, generally offers a better

option for safety studies. However, the design is also subject to selection bias capable of presenting similar problems. We suggest that by adopting the counter-matched version of the design and incorporating an appropriate group balancing instrument, we can reduce the problems of exposure misclassification and selection bias considerably to produce results robust enough to meet our objectives.

Of course, the cohort design offers the best option where the problem of intermittent drug exposure is not an issue and the case-cohort design can be a viable alternative to a full cohort. Indeed, the design can be utilized even where the problem exists. Although the case-cohort design is widely recognised as efficient and conceptually simple, it is rarely utilised in non-interventional PASS largely due to perceived analytical complexities. The design offers an option for continued collection of all relevant data on patients in a random subset of a registry that is about to come to its end. Indeed, where feasible, this design can complement current efforts aimed at improving the usefulness of Phase IV trials on drug safety. It offers a means for an extended follow-up of some of the study patients and for assessing multiple potential safety outcomes at improved statistical efficiency and without the huge resources which extended trials would entail. We see these as sufficient reasons for further assessment of the design and its consideration for application in post-market drug safety studies.

References

- 1 Miettinen OS and Caro JJ. Principles of nonexperimental assessment of excess risk, with special reference to adverse drug reactions. *J Clin Epidemiol* 1989; 42(4): 325-331.
- 2 Guess HA. Behavior of the exposure odds ratio in a case-control study when the hazard function is not constant over time. *J Clin Epidemiol* 1989; 42(12): 1179-1184.
- 3 Suissa S. Novel approaches to pharmacoepidemiology study design and statistical analysis. In: Strom BL, editor. *Pharmacoepidemiology*, 3rd ed. New York: John Wiley & Sons 2000; 785–805.
- 4 Black C, Kaye JA and Jick H. Relation of childhood gastrointestinal disorders to autism: nested case-control study using data from the UK general practice research database. *BMJ* 2002; 325 (7361): 419-421.
- 5 Eland IA, Huerta-Alvarez C, Stricker B, et al. The risk of acute pancreatitis associated with acid-suppressing drugs. *Br J Clin Pharmacol* 2000; 49: 473-478.
- 6 Nightingale AL and Farmer RD. Ischemic stroke in young women: a nested case-control study using the UK general practice research database. *Stroke* 2004; 35: 1574-1578.
- 7 Clayton D and Hill M. *Statistical Models in Epidemiology*. Oxford: Oxford University Press; 1993.
- 8 Greenland S and Neutra R. Control of confounding in the assessment of medical technology. *Int J Epidemiol*. 1980; 9(4): 361–367.
- 9 Horwitz RI, McFarlane MJ, Brennan TA and Feinstein AR. The role of susceptibility bias in epidemiologic research. *Arch Intern Med*. 1985; 145(5): 909-912.
- 10 Petri H and Urquhart J. Channelling bias in the interpretation of drug effects. *Stat Med* 1991; 10: 577-581.
- 11 Rosenbaum PR and Rubin DR. The central role of propensity score in observational studies for casual effects. *Biometrika* 1983; 70(1): 41-55.
- 12 D'Agostino RB. Tutorial in biostatistics- Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statist Med* 1998; 17: 2265-2281.

- 13 Morant SV, Pettitt D, MacDonald TM, et al. Application of propensity score to adjust for channelling bias with NSAIDs. *Pharmacoepidemiol Drug Saf* 2004; 13: 345-353.
- 14 Frolkis JP, Pthier CE, Blackstone EH, et al. Frequent ventricular ectopy after exercise as a predictor of death. *N Engl J Med* 2003; 348-349.
- 15 Heckman JJ. Sample Selection Bias as a Specification Error. *Econometrica* 1979; 47(1): 153-161.
- 16 Crown WH, Obenchain RL, Englehart L, et al. The application of sample selection models to outcomes research: the case of evaluating the effects of antidepressants therapy on resource utilization. *Stat Med* 1998; 17:1943-1958.
- 17 Sartori AE. An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions. *Political Analysis* 2003; 11: 111-138.
- 18 Christopher W and Mare RD. Models for Sample Selection Bias. *Annu Rev Sociol* 1992; 18: 327-350.
- 19 Maclure M. The case cross-over design: a method for studying transient effects on the risk of acute events. *Am J of Epidemiol* 1991; 133: 144–153.
- 20 Warner L, Macaluso M, Austin HD, et al. Application of the Case-Crossover Design to Reduce Unmeasured Confounding in Studies of Condom Effectiveness. *Am J of Epidemiol* 2005; 161: 765–773.
- 21 Hernandez-Diaz S, Hernán MA, Meyer K, et al. Case cross-over and case-time-control designs in birth defects epidemiology. *Am J of Epidemiol* 2003; 158: 385-391.
- 22 Marshall RJ and Jackson RT. Analysis of case-crossover designs. *Stat Med* 1993; 12: 2333-2341
- 23 Park T, Ki M and Yi SG. Statistical analysis of MMR vaccines adverse events on aseptic meningitis using the case cross-over design. *Stat med* 2004; 23(12): 1871-1883.
- 24 Suissa S. The case-time-control design. *Epidemiology* 1995; 6: 248-253.
- 25 Suissa S. The case-time-control design: further assumptions and conditions. *Epidemiology* 1998; 9: 441-445.
- 26 Greenland S. Confounding and exposure trends in case-crossover and case-time-control designs. *Epidemiology* 1996; 7: 231–239.
- 27 Donnan PT and Wang J. The case-crossover and case-time-control designs. *Pharmacoepidemiol drug saf* 2001; 10(3): 259 - 262.

- 28 Schneider M, Gange SJ, Margolick JB, et al. Application of case cross-over and case-time-control study designs in analysis of time varying predictors of T-cell Homeostasis Failure. *Annals of epidemiology* 2005; 15 (2): 137-144.
- 29 Feldmann U. Epidemiologic assessment of risks of adverse reactions associated with intermittent exposure. *Biometrics* 1993; 49: 419-428.
- 30 Farrington CP, Nash J, Miller E. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *Am J Epidemiol* 1996; 143: 1165-1173. [Erratum, *Am J Epidemiol* 1998; 147: 93.]
- 31 Farrington CP. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics* 1995; 51: 228-235.
- 32 Smeeth L, Thomas SL, Hall AJ, et al. Risk of Myocardial Infarction and Stroke after Acute Infection or Vaccination. *N Engl J Med* 2004; 351: 2611-2618.
- 33 Prentice RL. A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* 1986; 73:1-11
- 34 Barlow WE. Robust variance estimation for the case-cohort design. *Biometrics* 1994; 50: 1064-1072
- 35 Barlow WE, Ichikawa L, Rosner D, Izumu S. Analysis of case-cohort designs. *J Clin Epidemiol* 1999; 52(12): 1165-1172
- 36 Cox D. Partial likelihood. *Biometrika* 1975; 62: 265-276
- 37 Cox DR, Oakes D. Analysis of survival data. *Monographs on Statistics and Applied Probability*. New York: Chapman and Hall; 1984.
- 38 Besag J. Efficiency of Pseudolikelihood estimation for simple Gaussian fields. *Biometrika* 1977; 64: 616-618
- 39 Sato T. Estimation of a common risk ratio in stratified case-cohort studies. *Statistics in Medicine* 1992; 11: 1599-1605
- 40 Reid N, Crepeau H. Influence functions for proportional hazards regression. *Biometrika* 1985; 72: 1-9
- 41 Lin DY, Wei LJ. The robust inference for the Cox-proportional hazards model. *JASA* 1989; 84: 1074-1078