Practice of Epidemiology

Nurse Staffing Level and Nosocomial Infections: Empirical Evaluation of the Case-Crossover and Case-Time-Control Designs

Stéphane Hugonnet1, Andrés Villaveces2,3, and Didier Pittet1

1 Infection Control Program, Department of Internal Medicine, University of Geneva Hospitals, Geneva, Switzerland.
2 Quality of Care Unit, Medical Directorate, University of Geneva Hospitals, Geneva, Switzerland.
3 Current affiliation: Department of Epidemiology, University of North Carolina, Chapel Hill, NC.

Received for publication June 20, 2006; accepted for publication December 20, 2006.

The authors compared a case-crossover design, a case-time-control design, and a cohort design to evaluate the effect of nurse staffing level on the risk of nosocomial infections. They evaluated two strategies, conditional logistic regression and generalized estimating equation, to analyze the case-crossover study. The study was performed among critically ill patients in the medical intensive care unit of the University of Geneva Hospitals, Geneva, Switzerland. Of 366 patients who stayed more than 7 days in the intensive care unit between 1999 and 2002, 144 developed an infection. The main reasons for admission were infectious (35.3%), cardiovascular (32.5%), and pulmonary (19.7%) conditions. A comparison of the three study designs showed that lower nurse staffing was associated with an approximately 50% increased risk of nosocomial infections. All analyses yielded similar estimates, except that the point estimate obtained by the conditional logistic regression used in the case-crossover design was biased away from unity; the generalized estimating equation yielded unbiased results and is the most appropriate technique for case-crossover designs. The case-crossover methodology in hospital epidemiology is a promising alternative to traditional approaches, but selection of the referent periods is challenging.

cross infection; epidemiologic methods; personnel staffing and scheduling

Abbreviation: ICU, intensive care unit.

The case-crossover design can be seen as a variant of the traditional case-control study in which each case is its own control. The design was first described by Maclure (1) and has been quite extensively used, for example, to investigate the effect of air pollution on various health outcomes (2–5), immediate risk factors for myocardial infarction (1), injury prevention studies (6, 7), or the association between vaccines and adverse events (8). The literature is more sparse in the field of infectious diseases or hospital epidemiology, but case-crossover designs have been used to study risk factors for hemorrhagic fevers (9), triggers of needle stick injuries (10), a food-borne outbreak (11), and condom effectiveness (12).

The principle of the design is to compare the exposure during a window of time shortly before onset of disease with the exposure frequency during control or referent times in the same subject (13). It is well suited to assess the effect of transient exposure, or exposure with a transient effect, on acute outcomes and brings considerable advantages to avoid the always difficult choice of the control group and to control for time-independent confounding factors related to the patient, two aspects that are particularly relevant in hospital epidemiology. Furthermore, this approach is resource efficient, as there is no need to allocate time for data collection among controls.

However, case-crossover studies are not free of methodological difficulties, such as time trends of the exposure or lack of independency of the exposure within subjects. Similar to traditional case-control studies where controls are supposed to provide the distribution of the exposure in the
MATERIALS AND METHODS

Study design and population

The study was based on a cohort of patients admitted to the medical intensive care unit (ICU) of the University of Geneva Hospitals over a 4-year period. The methods and main results have been described elsewhere (26). We conducted prospective on-site surveillance of all nosocomial infections acquired in the ICU and included all patients with a stay of 2 days or more. Collected data included patients’ demographic characteristics, admission diagnosis, admission severity score, Charlson comorbidity index (27), daily exposure to invasive devices and selected drugs, daily nursing acuity score, daily number of hospitalized patients, and number of nurses present at work. ICU-acquired infections were diagnosed according to standard and validated case definitions (22, 28–30).

Design strategy

The aim of the present study was to empirically investigate whether the case-crossover strategy is a suitable study design to assess risk factors for nosocomial infections. In particular, we aimed to quantify the potential bias arising from the case-crossover design and to compare several analytical strategies.

The main exposure was the daily nurse/patient ratio. We compared point estimates obtained by analyzing the data as a case-crossover design with those obtained by using a case-time-control design and a cohort study. To make comparisons meaningful, all point estimates were derived from the same study population. Only the first infection was considered, and only those patients with a stay of 8 days or more in the ICU up to the first infection were included. The comparison group was patients who remained free of ICU-acquired infection with a stay of 8 days or more. The cutoff point of 8 days was chosen to ensure that the patient’s stay was long enough to include a case period and at least one control period (figure 1).

Case period

The hypothesis to be tested was that lower nurse staffing would increase the likelihood of infection. The delay (incubation period) between the exposure (lower staffing) and infection onset is unknown and probably varies according to the amount of exposure, the type of infection, the microorganisms involved, and the patient. However, by definition (28–30), the incubation period of nosocomial infections is not shorter than 48 hours. Consequently, we defined the case period as being the time elapsing from 2 to 4 days before infection onset (figure 1).

Control and referent periods

Several referent periods were selected among infected patients: 5–7 days, 8–10 days, and 11–13 days before infection (figure 1). Among noninfected patients, control periods were selected to mirror those of the infected patients: 2–4 days, 5–7 days, 8–10 days, and 11–13 days before discharge (figure 1). Patients could contribute to one or more control/referent periods, depending on their ICU length of stay.
Main exposure and covariates

The main exposure was the mean nurse/patient ratio over the different case, referent, and control periods. We then dichotomized the exposure using a nurse/patient ratio of 1.9 as a cutoff; values below 1.9 were considered as exposed. This cutoff point corresponds to the median nurse/patient ratio over the whole study period. Time-dependent covariates (such as the nurse/patient ratio or exposure to a central venous line) were measured during each period (case, referent, and control periods) and were consequently allowed to change with time. In addition, we allowed for a latency between exposure and infection.

Statistical analysis

We first graphically explored the nurse/patient ratio over the whole study period, assessed the stationarity and seldasticity of the time series (a time series is said to be stationary if its mean and variance are independent of time), and estimated the autocorrelation over time.

Case-crossover approach. In the case-crossover design, the nurse/patient ratio during the case window was contrasted with the ratio over the referent periods among infected patients. We used two different analytical methods. By analogy with traditional matched case-control studies, we first computed the odds ratio and 95 percent confidence intervals by conditional logistic regression. Because lack of independence and stationarity in the exposure series might invalidate the results obtained by conditional logistic regression, we then estimated the exposure odds ratio by use of generalized estimating equations (8, 31). We specified the outcome to follow a binomial distribution and the cluster unit to be the patient. We tested different correlation matrices that provided different results. Considering the shape of the autocorrelation, the model was run using an autoregressive matrix.

Case-time-control approach. First, only noninfected patients were considered. Exposure during the last control period (2–4 days before discharge) was contrasted with exposure during all other periods, and the odds ratio was obtained by conditional logistic regression. The exposure odds ratio among cases obtained in the case-crossover approach was then divided by the odds ratio obtained among noninfected patients to adjust for time trend.

Cohort approach. We finally used the entire cohort of infected and noninfected patients and the estimated odds ratio by generalized estimating equations. We estimated odds ratios instead of relative risks or rates so that eventual differences in point estimates compared with that of the case-crossover design would be explained mainly by different designs, rather than by different analytical strategies. In this design, we contrasted exposure during the case period (among infected patients) with exposure during all referent and control periods. Model specification was identical to that described in the case-crossover approach.

Adjustment for potential confounding factors was done as follows in the same way for all designs. In the princeps paper analyzing the full cohort (26), the following variables were considered in univariate analysis: age, gender, admission diagnosis, APACHE II (Acute Physiology and Chronic Health Evaluation II, a severity-of-disease classification system) score and nursing severity score at admission, co-morbidities, nurse/patient ratio, nurses’ training level, exposure to invasive devices (central and peripheral venous lines, peripheral arterial line, endotracheal tube, urinary catheter, nasogastric tube, drains), and exposure to selected drugs (prophylactic and therapeutic antibiotics, gastric antacids, total parenteral nutrition). Only variables associated with the outcome with a $p < 0.05$ were retained in the
multivariate model. Apart from the nurse/patient ratio, these variables were as follows: exposure to a central venous line, urinary catheter, endotracheal tube, and therapeutic antibiotics. In the present study, these same variables were used in all multivariate analyses to allow for meaningful comparison across designs.

RESULTS
Study population

Of 366 patients who stayed more than 7 days until discharge or infection in the ICU, 144 developed an infection. Characteristics of the study population are shown in table 1. The median age was 68 years (interquartile range: 56–76 years) and did not differ between infected and noninfected patients. The main reasons for admission were infectious, cardiovascular, and pulmonary conditions. The median admission APACHE II score was 28 (interquartile range: 22–35.5) and was significantly higher among infected (median = 30; interquartile range: 23–36) than noninfected (median = 27; interquartile range: 22–34) patients ($p = 0.049$). The median length of stay up to discharge or infection was 10 days (interquartile range: 9–12 days) and was similar among infected and noninfected patients. In total, there were 144 case periods, 237 referent periods (among the infected), and 570 control periods (among the noninfected) (tables 2 and 3).

Staffing level

The median daily nurse/patient ratio over the study period (1,224 days) was 1.9 (interquartile range: 1.8–2.2). The daily nurse/patient ratio was not independent over time, as the nurse/patient ratio values were correlated between successive days, for up to 13 days. The autocorrelation coefficient was 0.60 between 1-day lag observations and was 0.15 for a 13-day lag.

Association between staffing level and nosocomial infections

The crude and adjusted associations between staffing level and infection are shown in table 2 (case-crossover and cohort designs) and table 4 (case-time-control design). The nurse/patient ratio below 1.9 was associated with an increased infection risk in both case-crossover and cohort designs. However, the magnitude of the association differed according to the design and analytical strategy. In the case-crossover study, the point estimate obtained by conditional logistic regression was further away from unity than that obtained by generalized estimating equations and in the cohort study. On the other hand, exposure odds ratios obtained by generalized estimating equations, whether in the case-crossover or the cohort design, were similar (table 2).

Among noninfected patients, exposure to lower staffing was more frequent just before discharge; the adjusted exposure odds ratio among noninfected patients was 1.53 (95 percent confidence interval: 1.14, 2.06). Using this estimate in the case-time-control approach to remove the effect of time trend among cases gave an adjusted odds ratio of 1.24 in the case-crossover design (table 4), slightly lower than the estimate obtained in the cohort study.

DISCUSSION

This study confirms the association between staffing level and infection risk among critically ill patients (26), and it provides an empirical evaluation of the case-crossover and case-time-control designs in the field of hospital epidemiology. This is one of the very few studies in the field of infectious diseases or hospital epidemiology that uses the case-crossover design (9, 11, 12), compares the case-crossover
approach with other study designs, and evaluates several statistical techniques on real data (9, 12).

By design, case-crossover studies automatically control for time-independent patient confounding factors, since case and referent periods belong to the same patients. This results in a perfect match on known, but more importantly unknown, risk or confounding factors and is particularly relevant in hospital epidemiology. Risk factors for infection are only partially understood, and what is known about risk factors explains only a relatively small proportion of all infections. There are numerous confounding factors related to patients. Some are usually measured, such as age, severity of illness at admission, comorbidities, and admission diagnosis, but many are either unknown and/or unmeasured. Warner et al. (12) investigated the association between condom use and incident gonorrheal and chlamydial infections by a case-crossover analysis and a cohort analysis. The results differed considerably, as condom use had no effect in the cohort analysis but was associated with a reduced incidence in the case-crossover analysis. This difference was attributed by the authors to adjustment for unmeasured confounding factors.

The selection of the referent periods is a crucial step in case-crossover studies and poses difficult challenges that are no easier to solve than the selection of control patients in traditional case-control studies. One difficulty is related to the characteristics of the exposure series in terms of time trend in the exposure or lack of independency between observations. If the exposure series shows a time trend, selecting referent periods before the case period will lead to a systematic selection bias of the referent periods. The exposure odds ratio will be biased, and the direction of the bias will depend on the type of time trend. This is very relevant when studying the effect of air pollution, as its level is dependent on the season and day of the week. The way referent periods are selected can adjust for the time trend in the exposure, and several strategies have been proposed and used. One is to sample several referent periods before the case period, with an interval between them that is a function of the trend (10, 32, 33). Another strategy is to select referent periods before and after the outcome of interest, that is, bidirectional sampling (2, 14, 34). None of these strategies can be used when investigating risk factors for nosocomial infection in critically ill patients. First, the length of stay is short and surely shorter than possible trends in any exposure, thus leaving no margin for the selection of the referent period, neither in terms of number of referent periods nor in terms of timing compared with the case period. Second, selecting the referent period after the outcome can be done only if the outcome does not affect the exposure. This is true in air pollution studies, since being hospitalized as a consequence of air pollution does not impact on the level of air pollution. This will, however, be wrong most of the time in the field of nosocomial infections. In the present study, the exposure of interest was the nurse/patient ratio, a surrogate marker of workload. The occurrence of an infection will result, for instance, in additional care for the patient, risk of re-infection, or the hospitalization of the patient at another hospital.


<table>
<thead>
<tr>
<th></th>
<th>No. of patients</th>
<th>No. of periods</th>
<th>Nurse/patient ratio of &lt;1.9*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Infected</td>
<td>Noninfected</td>
<td>Case period</td>
</tr>
<tr>
<td>Case-crossover</td>
<td>144</td>
<td>0</td>
<td>144</td>
</tr>
<tr>
<td>Conditional</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>logistic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalized</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>estimating</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort design</td>
<td>144</td>
<td>222</td>
<td>144</td>
</tr>
<tr>
<td>(generalized</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>estimating</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Referent category is a nurse/patient ratio of ≥1.9.
† Adjusted for central venous line, mechanical ventilation, urinary catheter, and therapeutic antibiotics.

### TABLE 3. Number of patients and number of case periods and control periods in the case-time-control design, University of Geneva Hospitals, Geneva, Switzerland, 1999–2002

<table>
<thead>
<tr>
<th></th>
<th>Infected patients</th>
<th>Noninfected patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of patients</td>
<td>144</td>
<td>222</td>
</tr>
<tr>
<td>Infected patients</td>
<td>144</td>
<td>0</td>
</tr>
<tr>
<td>Noninfected</td>
<td>0</td>
<td>222</td>
</tr>
<tr>
<td>No. of periods</td>
<td>381</td>
<td>570</td>
</tr>
<tr>
<td>Case period</td>
<td>144</td>
<td>222</td>
</tr>
<tr>
<td>Control period</td>
<td>237</td>
<td>348</td>
</tr>
</tbody>
</table>

* The case period among noninfected patients corresponds to the time just before discharge.
insertion of a vascular line, administration of antibiotics, and, in other words, an increased workload.

The case-time-control design has been proposed to account for exposure time trends (15, 16) and has been used, for instance, in birth defect epidemiology (17). This design requires information on infected patients as well as on controls. The time trend in the nurse/patient ratio is provided by the odds ratio among controls, and this is used to adjust the exposure odds ratio among cases. Interestingly, the odds ratio obtained by the case-time-control design is close to, although lower than, that obtained in the cohort analysis or the case-crossover design analyzed by use of generalized estimating equations. However, we doubt that the case-time-control design is applicable in this situation. Adjusting for the time trend by use of the exposure odds ratio among controls implies that the pattern of the time trend somehow parallels that of the cases. Fluctuation in the nurse/patient ratio is explained mainly by the variation in the number of patients; moreover, the number of nurses on a given day is not imposed by anticipated discharges, and patients are not discharged earlier than they should be because of nurse shortage. Consequently, as controls were not matched to cases on time, there is no reason why the pattern of the exposure time trend should be similar among controls. If we had matched controls to cases on time, then the time trend would be strictly similar, as would be the exposure odds ratios among cases and controls, and this would yield an adjusted odds ratio of one.

The exposure during the referent period needs to be representative of the entire cohort for the result to be valid, just as in traditional case-control studies where the exposure among controls should be representative of the base population. In our study, the staffing level during the referent periods is representative of the entire cohort and hence, as shown, our estimates are similar for both the case-crossover and cohort designs. This would not necessarily be the case if the exposure under study were confounded by extrinsic factors.

The most appropriate statistical approach by which to analyze case-crossover studies remains a subject of debate, and much of the discussion is based on theoretical arguments or simulation studies (4, 31, 35, 36). It has been argued that using conditional logistic regression in the case-crossover design might provide biased estimates when the exposure series is correlated or nonstationary. The estimate that we obtained from the case-crossover design by conditional logistic regression was over 30 percent higher than that obtained with generalized estimating equations. The exposure series showed a clear time trend with a strong auto-correlation. Generalized estimating equations accounted for this trend by specifying an autoregressive correlation structure.

A carryover effect occurs when the effect of the exposure during a referent period lasts long enough to have an influence on the occurrence of the outcome. Indeed, the staffing level was higher during referent periods. If a higher staffing level had a long enough lasting effect that could influence (decrease) the infection risk, then the relation between a low staffing level during case periods and infection would be diluted and thus biased toward the null. Removing the carryover effect can be achieved by spacing the referent period from the case period. This was problematic in our situation, given the short length of stay of some of the patients.

Our study suffers some limitations. First, our case-crossover design is unable to control for overlap bias, given the absence of an appropriate time-stratified selection scheme for our referent periods (5). Overlap bias occurs when the effect of an exposure measured during the referent period develops with some delay, thus influencing the outcome; this bias is usually small in magnitude. However, given the strength of the association, it is unlikely that this bias would substantially change the results. Furthermore, if present, it would probably bias the result toward the null. Second, only patients staying long enough (at least 8 days in the present study) were included; consequently, results cannot be inferred to patients staying less than 8 days, and such a selection reduces the sample size and power of the study.

The case-crossover design is an interesting strategy in that it controls by design for some confounding, bypasses the need for a control group, and decreases the time required for data collection. However, the selection of the referent
periods is challenging. Analysis of the case-crossover design by generalized estimating equations provided unbiased results and is superior to conditional logistic regression. Using this methodology to study health-care–associated infections is a promising alternative that should be tested further against traditional approaches.

ACKNOWLEDGMENTS

The study was supported by research grant 32-68164.02 from the Swiss National Science Foundation.

The authors thank Nadia Colaizzi for data management and Rosemary Sudan for providing editorial assistance. They also thank Dr. Carmen Lucia Pessoa-Silva, Dr. Hugo Sax, Dr. Stephan Harbarth, and Dr. Thomas Permeier for their advice regarding the statistical analysis and review of the manuscript.

Conflict of interest: none declared.

REFERENCES