Papers

Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals

Philippe Michel, Jean Luc Quenon, Anne Marie de Sarasqueta, Olivier Scemama

Abstract

Objectives To compare the effectiveness, reliability, and acceptability of estimating rates of adverse events and rates of preventable adverse events using three methods: cross sectional (data gathered in one day), prospective (data gathered during hospital stay), and retrospective (review of medical records).
Design Independent assessment of three methods applied to one sample.
Setting 37 wards in seven hospitals (three public, four private) in southwestern France.
Participants 778 patients: medical (n = 278), surgical (n = 263), and obstetric (n = 237).
Main outcome measures The main outcome measures were the proportion of cases (patients with at least one adverse event) identified by each method compared with a reference list of cases confirmed by ward staff and the proportion of preventable cases (patients with at least one preventable adverse event). Secondary outcome measures were inter-rater reliability of screening and identification, perceived workload, and face validity of results.
Results The prospective and retrospective methods identified similar numbers of medical and surgical cases (70% and 66% of the total, respectively) but the prospective method identified more preventable cases (64% and 40%, respectively), had good reliability for identification ($\kappa = 0.83$), represented an acceptable workload, and had higher face validity. The cross sectional method showed a large number of false positives and identified none of the most serious adverse events. None of the methods was appropriate for obstetrics.
Conclusion The prospective method of data collection may be more appropriate for epidemiological studies that aim to convince clinical teams that their errors contribute significantly to adverse events, to study organisational and human factors, and to assess the impact of risk reduction programmes.

Introduction

Review of medical records is considered the benchmark for estimating the extent of medical injuries in hospitals.[1–7] But the limitations of this retrospective method raise the issue of alternative methods, especially since the epidemiology of adverse events and medical errors is moving towards other objectives, such as assessing the impact of risk reduction programmes and studying organisational and human factors.[8–10]

We compared prospective and cross sectional methods for data collection with review of medical records for assessing rates of adverse events and rates of preventable adverse events in acute care hospitals in France. Although the cross sectional method is used to estimate prevalence and not incidence, we chose to study it because in France it is usually used to assess certain risks such as care related infections or adverse drug reactions.[11 12]

Methods

An adverse event was defined as an unintended injury caused by medical management rather than by a disease process and which resulted in death, life threatening illness, disability at time of discharge, admission to hospital, or prolongation of hospital stay.[2–5] Preventable adverse events were those that would not have occurred if the patient had received ordinary standards of care appropriate for the time of the study.

Our population sample was inpatients in medical, surgical, and obstetric wards in acute care hospitals in Aquitaine, southwestern France (3 million inhabitants and 14 000 hospital beds). The sample was obtained by a two stage cluster stratified process (see figure on bmj.com). For each hospital in each stratum, we selected wards proportional to the number of hospital beds (proportional allocation).

Independent investigators consecutively applied the three methods to the sample (fig 1). We used two questionnaires, which were adapted from an English survey.[5] One was used for detection (by nurses and midwives) and one for confirmation (by doctors). The questionnaires were used three times for each patient, one for each method. The detection and confirmation questionnaires used for record review and the questionnaire used for data collection from clinical teams were similar. Possible adverse events were detected by two nurses in medicine and surgery and two midwives in obstetrics. One nurse or midwife performed the cross sectional and prospective methods and the other reviewed the medical records. Three fully qualified doctors in each ward identified cases, one for each method. Each doctor participated equally in the three methods.

Data collection

Cross sectional method

Patients were included on the day that the cross sectional method was performed, when the nurses or midwives interviewed the head nurse and if necessary consulted the patient's medical records on the basis of 17 criteria (box 1). Patients who screened positive were referred to the physician investigator, who interviewed the doctor (resident or senior doctor) who managed the patient on the day of data collection,

P+ *The sampling strategy is on bmj.com*

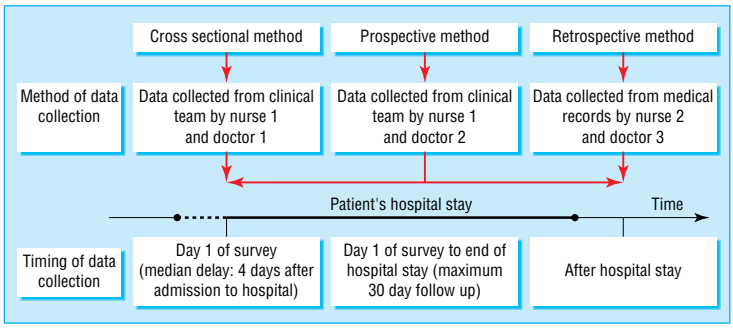

Fig 1 Method and timing of data collection

and consulted the patient's medical records if necessary. That day constituted the first day of the prospective method.

Prospective method

For the prospective method the detection investigators visited the ward on day one of the survey and on two other occasions during the first seven days, then once a week for up to a month. The doctor involved in the prospective method visited the ward at the end of the first week then when the last patient was discharged or on day 30 if patients were still present. Thus patients with adverse events detected on the first day were confirmed by two different doctors one week apart.

Retrospective method

For the retrospective method, review of the medical records began 30 days after the cross sectional method. The proportion of cases (incidence rate for retrospective and prospective methods, point prevalence for cross sectional method) was computed by taking into account the hospital's cluster effect, using the svy program (release 5.0; STATA).

Box 1: Criteria used when consulting medical records

- Unplanned admission as a result of any healthcare management during the 12 months before index admission
- Hospital incurred unintentional patient injury
- Adverse drug reaction or drug error
- Hospital acquired infection or sepsis
- Unplanned removal, injury, or repair of organ or structure during surgery, invasive procedure, or vaginal delivery
- Unplanned return or visit to the operating theatre
- Unplanned open surgery after closed or laparoscopic surgery
- Cardiac or respiratory arrest, low Apgar score
- Development of neurological deficit not present on admission
- Injury or complications related to termination of labour, and delivery including neonatal complications
- Other patient complications including myocardial infarction, deep vein thrombosis, cerebrovascular accident, pulmonary embolism
- Patient or family dissatisfaction with care received documented in the medical record, or documentation of claim or litigation
- Unplanned transfer from general care to intensive care or higher dependency
- Unplanned transfer to another acute care hospital
- Unexpected death (that is, not an expected outcome of the disease during hospital stay)
- Documented pain or psychological or social injury
- Any other undesirable outcomes (not covered by any of the other criteria)

Outcome measures and statistical analysis

Effectiveness of the methods was determined by the proportion of cases identified in relation to a reference list. This list was based on the adverse events identified by any one of the three methods. We checked each adverse event with the ward doctors and with the investigators to resolve any conflicts. We calculated the proportion of preventable cases (at least one preventable adverse event) detected by each method in relation to the reference method. We used paired χ^2 tests to compare the effectiveness of the prospective and retrospective methods for identifying cases, preventable cases, and their subgroups (adverse events occurring during hospital stay, those responsible for all or part of hospital admission, and the most serious—associated with death or life threatening illness). For these subgroup analyses, we aggregated data from the medical and surgical wards using an equivalent weight, because in France the number of beds in these specialties is similar.[13]

Reliability of the prospective method was assessed from cases detected on the first day of the study and confirmed twice—by the doctors performing the cross sectional and prospective methods. Acceptability of the workload and face validity of the results were assessed during sessions with each clinical team after data analysis in the clinical wards. Questions were open ended and participants could express their perception of the workload and of the truthfulness of the results of each method.

Results

We selected 37 wards in three public and four private hospitals: medical (n = 15), surgical (n = 14), and obstetric (n = 8). Overall, 786 patients were included on the day of the cross sectional method. Eight were excluded because they were still present on day 30, precluding the review of their medical records. The three methods were therefore applied to 778 patients: 278 in medicine, 263 in surgery, and 237 in obstetrics. The adverse event rates found by the prospective and retrospective methods were similar and the point prevalence obtained by the cross sectional method was about one third lower than that of the other two methods (table). The incidence of preventable adverse events assessed prospectively was 25% higher than that assessed retrospectively.

Thirteen of the 254 adverse events were excluded from the reference list and considered as false positives (12 of the adverse events identified by the cross sectional method, none by the prospective method, and one by the retrospective method). The final list comprised 241 adverse events in 174 patients (110/80 in medicine, 114/80 in surgery, and 17/14 in obstetrics). Of these 174 patients, 122 (70%) had one adverse event, 38 (22%) had two adverse events, and 14 (8%) had three adverse events (125 patients had at least one adverse event on the day of cross

Proportion of inpatients with at least one adverse event according to epidemiological method (n=778)

	Retrospective		Prospective		Cross sectional	
Characteristic	**No**	**% (95% CI)***	**No**	**% (95% CI)**	**No**	**% (95% CI)**
Patient with adverse event	112	14.5 (10.4 to 18.7)	120	15.4 (12.2 to 8.7)	76	9.8 (6.8 to 12.8)
Patient with preventable adverse event	31	4.0 (5.2 to 6.7)	50	6.4 (4.9 to 8.0)	27	3.5 (2.0 to 4.9)

*Hospital's cluster effect taken into account.

sectional data collection). The cross sectional method identified 39 (64%) of the medical cases, 32 (56%) of the surgical cases, and 5 (45%) of the obstetric cases, and, respectively, 18 (51%), 6 (27%), and 3 (43%) of the preventable ones (fig 2). None of the most serious adverse events were identified.

The prospective method identified 63 (79%) of the medical cases, 49 (61%) of the surgical cases, and 8 (57%) of the obstetric cases, and, respectively, 32 (74%), 12 (46%), and 4 (44%) of the preventable cases. The retrospective method identified 43 (54%) of the medical cases, 61 (76%) of the surgical cases, and 8 (57%) of the obstetric cases, and, respectively, 16 (37%), 12 (46%), and 3 (33%) of the preventable cases. The prospective method was significantly more effective than the retrospective method at identifying cases in medicine (paired $\chi^2=8.64$, $P<0.01$) and tended to be less effective at identifying cases in surgery (paired $\chi^2=3.24$, $P<0.07$). The prospective method was significantly more effective at identifying preventable cases in medicine (paired $\chi^2=3.0$, $P<0.005$) but no difference was observed in surgery. In obstetrics, effectiveness was not different but there were too few adverse events to draw conclusions.

When the medical and surgical cases were aggregated, the prospective and retrospective methods showed similar effectiveness (70% and 66% of 160 patients identified, respectively), but the prospective method was more effective at identifying preventable cases (64% and 40% of 71 patients, respectively; $P<0.02$). Effectiveness was similar for adverse events occurring during hospital stay (68% and 64% of 110 patients), those responsible for hospital admissions (63% and 59% of 75 patients), and the most serious events (44% and 33% of 52 patients).

Reliability of cases identified among the 145 cases detected on the first day of study was good (global agreement 91.7%; $\kappa=0.83$, 95% confidence interval 0.67 to 0.99), but agreement about preventability was low (67.8%; $\kappa=0.31$, 0.05 to 0.57).

The workload for the prospective and cross sectional methods was perceived as similar, as the phase considered most time consuming for the staff was detection on the first day of data collection (average three hours in wards with 25 patients). Workload was perceived as less for the retrospective method but not negligible, especially when the information sources were multiple and the search was performed by the secretarial staff. The hospital staff constantly preferred the cross sectional and prospective methods because of their pedagogical and communicative virtues.

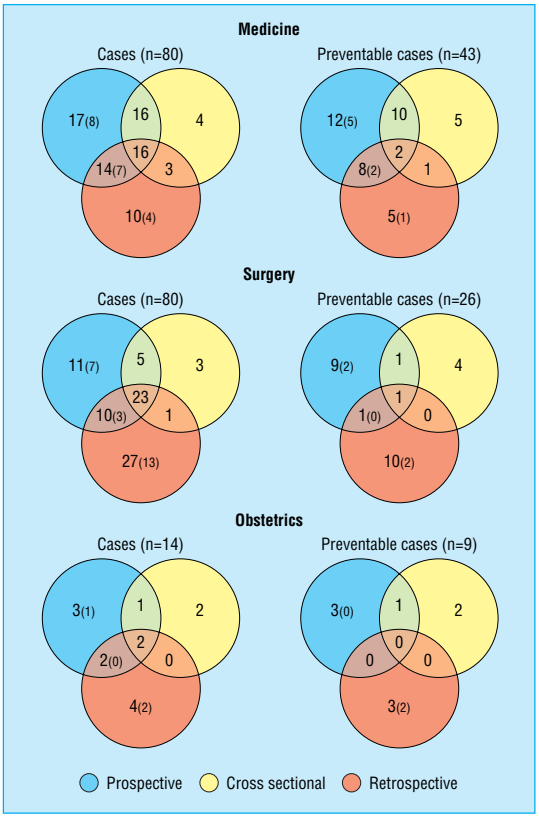

Fig 2 Venn diagrams showing number of cases identified by three methods of data collection. Cases in brackets not identified by cross sectional method as they were identified after first day of data collection

Discussion

The retrospective method of data collection by review of medical records is as effective as the prospective method (data gathered during hospital stay) for estimating adverse event rates. Given the presently evolving aims of epidemiology for adverse events, the prospective method has several advantages over retrospective and cross sectional (data gathered on a given day) methods. It identified more preventable cases, was more reliable, had better face validity, and involved an acceptable workload.

Limitations of study

The external validity of perception of workload is questionable because it was assessed in a limited number of care teams without using a structured interview. The external validity of effectiveness and reliability seems reasonable because our sample was obtained by random sampling and stratification and because we applied the three methods to the same sample. Nevertheless, bias may have been present due to the small number of hospitals and wards. In medicine especially, the preventable cases may have been under-represented because two of the 15 wards specialised in oncology and accounted for 23 of the 110 (21%) adverse events, of which only five were preventable. Incidence was not an objective of our study and is subject to biases from inclusion of patients on a given day and because we did not take into account adverse events between admission to hospital and the first day of data collection (median four days). The incidence of adverse events found by review of the medical records was, however, consistent with previous findings.[2 5 7 14] Another limitation concerns the reference list. Because there is no gold standard for the reference list, we checked all cases with the investigators and the clinical teams to resolve conflicts.[14] The list may have contained errors because it was established by doctors, and not all the doctors involved in adverse events participated in this process. However, the validity

of the list would have been questionable if doctors had agreed with cases identified by the prospective method (because they were involved in their identification) and had invalidated many of the cases identified by the retrospective method; only one case was invalidated. Another drawback to internal validity was better than usual identification of adverse events by the retrospective method because the care teams were involved in the identification of events during the prospective method. This bias in favour of the retrospective method was not likely since the number of cases identified by the retrospective or prospective methods was similar. All eight patients who were still in the clinical wards at the beginning of the retrospective method (so medical record review could not be performed) had adverse events according to the prospective method. This may have resulted in bias in favour of the prospective method.

Box 2: Advantages and disadvantages of three methods used to estimate adverse event rates

Prospective method (data collected during hospital stay)
Advantages
- Best effectiveness for identifying preventable events
- Good reliability of judgment of iatrogenic nature of events
- Staff sufficiently involved to understand notion of iatrogenic risk and search for causes
- Preferred because of their pedagogical and communicative virtues
- Good appreciation of chain of events and their consequences
- Possible role as "red flag" for care providers during data collection

Disadvantages
- Most expensive
- Heaviest workload, although perceived as acceptable:
  Several visits for investigators
  Staff must be available

Cross sectional method (data gathered on given day)
Advantages
- Least expensive
- Seamless continuation of former methodological approaches to iatrogenic risk
- Methodological approach fully understood by professionals and appreciated because it is rapid and easily renewed
- May be sufficient to justify implementation of a risk reduction policy and to define priorities
- Good reliability of judgment of iatrogenic nature of events; possible role as "red flag" for care providers during data collection

Disadvantages
- Consequences of lack of follow up during patient's hospital stay:
  Lowest effectiveness
  Lack of validity due to measurement errors (false positives and false negatives)
  Prevalence biased by underestimation of frequency, particularly of deaths, and by over-representation of short stays
- Believed by unit staff to involve an excessive workload for obtaining imprecise estimations
- Inadequate to serve as initial estimation when evaluating impact of risk reduction policy

Retrospective (review of medical records)
Advantages
- Good effectiveness, even superior in surgery for estimating adverse event incidence
- Almost no workload for staff
- Data collection easily planned
- Method sometimes favoured by surgical teams and centres

Disadvantages
- Difficulty to judge iatrogenic and preventable nature on basis of sometimes piecemeal data. Therefore:
  Measurement errors due to quality of medical records and to lesser reliability of judgment of iatrogenic nature
  Underestimation of preventable events
- Lower face validity of results, especially for preventability judgment (no involvement of staff)

Strengths of study

One strength of our study was the different results shown according to ward type by stratification. Unlike our results, another study found that the prospective method detected twice to four times more complication rates in surgical patients than did a review of medical records.[15] These results may be due to differences in definitions or in data collection. The prospective method has proved more effective than the retrospective method for specific risks.[16 17]

In contrast to medicine and surgery, data on the incidence of obstetrical adverse events are sparse.[18] Most studies focus on a limited number of events.[19] The performance of methods and the definition of adverse events that led to a high frequency of positive screening and low frequency of identification should be questioned. Adverse events in obstetrics are, however, of major importance even at low rates, because the patients are generally healthy women. Large studies are needed to determine whether lowering the level of severity of adverse events or pursuing follow up systematically after discharge would adequately increase the performance of the epidemiological methods. We did not include obstetric cases in our aggregated analyses because of their rarity and clinical difference to medical and surgical ones.

Prevalence surveys are commonly conducted for specific risks.[12 20 21] Such methods may be more valid for adverse drug reactions that occur before an index admission and that were the reason for admission.[10] Overall, 11 of the 28 nosocomial infections in our study that should have been identified by the cross sectional method were missed. Such false negatives may be an issue in epidemiological studies that follow trends over time. The cross sectional method should be used cautiously to assess the incidence of adverse events globally because of the high rate of false positives and false negatives. Because we expected the cross sectional method to be less effective, we aimed to assess the extent of the lack of effectiveness. The cross sectional method is less expensive than the retrospective or prospective methods, is easier for non-epidemiologists to understand, and allows clinical team to collect data.

Conclusions

Our results provide new insights into the epidemiology of adverse events. Firstly, they suggest ways to improve prospective assessment. The under identification of cases was mostly due to medical staff inadequately grasping the concept of adverse events. For example, they did not consider events as adverse if they were frequent or the patient recovered without sequelae. Under identification could be reduced by providing doctors with detailed information before data collection and by reviewing medical records before interviewing the doctors. Secondly, our results suggest that the prospective method is preferable for assessing the impact of risk reduction programmes (better reliability), for convincing clinical teams that their errors may contribute significantly to adverse events (better face validity resulting partly from team's involvement in data collection), for improving the assessment of consequences, and for studying

What is already known on this topic

Estimates of adverse event rates from large studies are based on review of medical records, despite limitations

The epidemiology of adverse events is moving towards more analytical and evaluative goals for which these limitations may be serious issues

What this study adds

No reference method exists for identifying adverse events

The retrospective method is appropriate for estimating rates of adverse events

The prospective method, based on data gathered from wards, should be preferred for describing causes and consequences of adverse events and for evaluating risk reduction programmes

The type of epidemiological method should be chosen according to objectives

organisational and human mishaps (especially better effectiveness in identifying preventable adverse events).

We thank for their advice Charles Vincent, Maria Woloshynowych, and Graham Neale; Dominique Baubeau, Anne Broyard-Farge, Chantal Cases, Mireille Elbaum, Brigitte Haury, Jacques Massol, and Florence Veber (Ministry of Health); Pascal Astagneau, Françoise Haramburu, Lionel Pazart, and Marie-Laure Pibarot (national working group); the participating hospitals (university hospital of Bordeaux, Libourne and Langon public hospitals, Saint Augustin, Bordeaux Nord, Bordeaux Caudéran and Cenon Rive Droite private hospitals); the investigators, Françoise Zaro-Goni and Xavier Pelloquin (nurses), Delphine Caute and Nathalie Besson (midwives), Danielle Dreyfus, Sandrine Harston, and Michel Marcos (medicine), Alain Chabrol, Jean Pierre Claverie, and Astrid Sieber-Roth (surgery), and Jeanine Schirumberro, Philippe Cormier, and Richard Torrieli (obstetrics).

Contributors: PM designed the project, wrote the original research proposal, and managed the project. PM and JLQ supervised the physician investigators and performed data analysis. AMS supervised the nurse and midwife investigators and was responsible for data quality control. OS did the initial literature analysis and translated the English review form into our questionnaires. PM and JLQ will act as guarantors for the paper.

Funding: Financial support was provided by the Ministry of Health (Direction de la Recherche, des Etudes, de l'Evaluation et des Statistiques).

Competing interests: None declared.

Ethical approval: Not required.

1 Weingart SN, Wilson RM, Gibberd RW, Harrison B. Epidemiology of medical error. *BMJ* 2000;320:774-7.
2 Wilson RM, Runciman WB, Gibberd RW, Harrisson BT, Newby L, Hamilton JD. The quality in Australian Health-Care Study. *Med J Aust* 1995;163:458-71.
3 Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000;38:261-71.
4 Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG. Incidence of adverse events and negligence care in hospitalized patients: results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:370-6.
5 Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.
6 Schioler T, Lipczak H, Pedersen BL, Mogensen TS, Bech KB, Stockmarr A, et al. [Incidence of adverse events in hospitals. A retrospective study of medical records]. *Ugeskr Laeger* 2001;163:5370-8.
7 Davis P, Lay-Yee R, Schug S, Briant R, Scott A, Johnson S, et al. Adverse events regional feasibility study: indicative findings. *NZ Med J* 2001;114:203-5.
8 Neale G, Woloshynowych M. Retrospective case record review: a blunt instrument that needs sharpening. *Qual Safety Health Care* 2003;12:2-3.
9 Thomas EJ, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. *Ann Intern Med* 2002;136:812-6.
10 Weingart SN, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Discrepancies between explicit and implicit review: physician and nurse assessments of complications and quality. *Health Serv Res* 2002;37:483-98.
11 Pouyanne P, Haramburu F, Imbs JL, Bégaud B. Admissions to hospital caused by adverse drug reactions: a cross-sectional incidence study. French pharmacovigilance centres. *BMJ* 2000;320:1036.
12 The French Prevalence Survey Study Group. Prevalence of nosocomial infections in France: results of the nationwide survey in 1996. *J Hosp Infect* 2000;46:186-93.
13 Directions Régionales des Affaires Sanitaires et Sociales. *STATISS 2000. Les régions françaises*. Paris: Ministère de l'Emploi et de la Solidarité, 2000.
14 Brennan TA, Localio AR, Laird NM. Reliability and validity of judgments concerning adverse events and negligence. *Med Care* 1989;27:1148-58.
15 Healey MA, Shackford SR, Osler TM, Rogers FB, Burns E. Complications in surgical patients. *Arch Surg* 2002;137:611-7.
16 Glenister HM, Taylor LJ, Bartlett CLR, Cooke EM, Sedgwick JA, Mackintosh CA. An evaluation of surveillance methods for detecting infections in hospitals inpatients. *J Hosp Infect* 1993;23:229-42.
17 Flynn EA, Barker KN, Pepper GA, Bates DW, Mikeal RL. Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. *Am J Health-Syst Pharm* 2002;59:436-46.
18 Mantel GD, Biuchmann E, Rees H, Pattinson RC. Severe acute maternal morbidity: a pilot study of a definition for a near-miss. *Br J Obstet Gynaecol* 1998;105:985-90.
19 Waterstone M, Bewley S, Wolfe C. Incidence and predictors of severe obstetric morbidity: case control study. *BMJ* 2001;322:1089-93.
20 Grohskopf LA, Sinkowitz-Cochran RL, Garrett DO, Sohn AH, Levine GL, Siegel JD, et al. A national point-prevalence survey of pediatric intensive care unit-acquired infections in the United States. *J Pediatr* 2002;140:432-8.
21 Emmerson AM, Enstone JE, Griffin M, Kelsey MC, Smyth ET. The Second National Prevalence Survey of infection in hospitals—overview of the results. *J Hosp Infect* 1996;32:175-90.

(Accepted 16 October 2003)

bmj.com 2004;328:199

Comité de Coordination de l'Evaluation Clinique et de la Qualité en Aquitaine, Hôpital Xavier Arnozan, 33604 Pessac, France
Philippe Michel *medical director*
Jean Luc Quenon *epidemiologist*
Anne Marie de Sarasqueta *public health nurse*
Olivier Scemama *epidemiologist*
Correspondence to: P Michel
philippe.michel@ccecqa.asso.fr