Journal of Public Health | Vol. 35, No. 1, pp. 147–156 | doi:10.1093/pubmed/fds057 | Advance Access Publication 24 July 2012

Advantages and limitations of using national administrative data on obstetric blood transfusions to estimate the frequency of obstetric hemorrhages

C. Quantin[1,2], E. Benzenine[1], C. Ferdynus[3], M. Sediki[4], B. Auverlot[1], M. Abrahamowicz[5], P. Morel[6], J. B. Gouyon[3,7], P. Sagot[3,8]

[1]CHRU, Service de Biostatistique et d'Informatique Médicale, BP 77908, CHU de Dijon, Dijon 21079, France
[2]Inserm, U866, Univ de Bourgogne, Dijon F-21000, France
[3]Centre d'épidémiologie des populations, EA 4184, Université de Bourgogne, Dijon, France
[4]INSERM CIC-P 803, CHU de Dijon, Dijon, France
[5]Department of Epidemiology and Biostatistics, McGill University, Montreal, QC, Canada H3A 1A1
[6]Etablissement Français du Sang – Bourgogne-Franche-Comté, Besançon F-25020, France
[7]Faculté de Médecine, Université de Bourgogne, Dijon, France
[8]Service de gynécologie-obstétrique, Médecine fœtale et reproduction humaine, CHU de Dijon, Dijon F-21000, France
Address correspondence to C. Quantin, E-mail: catherine.quantin@chu-dijon.fr

ABSTRACT

Background Obstetric hemorrhages are a frequent cause of maternal death all over the world, but are not routinely monitored. Health systems administrative databases could be used for this purpose, but data quality needs to be assessed.

Objectives Using blood transfusion data recorded in administrative databases to estimate the frequency of obstetric hemorrhages.

Research design A population-based study.

Subjects Validation sub-sample: all mothers who gave birth in a French region in 2006–07 (35 123 pregnancies). Main study: all mothers who gave birth in France in 2006–07 (1 629 537 pregnancies).

Method Linkage and comparison of administrative data on blood transfusions with data from the French blood agency ('gold standard'), and, based on this validation, the construction of a multivariable regression model to correct the number of pregnant women identified as having received a transfusion in the national administrative database.

Results The blood transfusion rate observed in the gold standard was 7.12‰. The sensitivity of the administrative data was estimated at 66.3% and the positive predictive value at 91.3%. The estimated total number of pregnant women who received blood transfusions in France in 2006–07 was 10 941 (6.71‰).

Conclusions The administrative data, available in most countries, can be used to estimate the frequency of obstetric hemorrhages.

Keywords anonymization, blood cells unit, database, evaluation, linkage, maternal morbidity, *post partum* haemorrhage, perinatal network, transfusion

Introduction

Improving maternal health is the fifth of the eight Millennium Development Goals set by the United Nations in September 2000.[1] To assess this improvement in high-income countries, monitoring severe acute maternal morbidities[2–6] seems to be preferable to assessing the maternal mortality ratio (number of maternal deaths per 100 000 live

C. Quantin, Professor of Public Health
E. Benzenine, Research Assistant
C. Ferdynus, Research Assistant
M. Sediki, Research Assistant
B. Auverlot, MD specialized in Public Health
M. Abrahamowicz, Professor of Biostatistics
P. Morel, MD specialized in Hematology
J. B. Gouyon, Professor of Pediatrics
P. Sagot, Professor of Gynecology

births),[7–9] which is quite low in these countries, around 10/100 000[10,11] with only ~27% of deaths indirectly related to the pregnancy.[4,12] However, there are few nationwide studies of severe acute maternal morbidities,[13] and most are based on retrospective surveys,[14–18] rather than on prospective studies.[19] Monitoring severe acute maternal morbidities thus appears to be an innovative and effective way to assess the quality of maternal health care.

Major obstetric hemorrhage (MOH) is one of the most frequent causes of maternal morbidity worldwide.[9] Whereas there is no general consensus regarding the standard definition of major hemorrhages, there appears to be some agreement regarding the way MOH can be assessed in clinical practice: for example, the Healthcare Commission of the UK defined 'significant' blood loss as >1000 ml and 'major' blood loss as >2500 ml, in its recent review of maternity services in England and Wales.[20] However, it should be noted that blood loss, and in particular maternal blood loss, is notoriously difficult to quantify and tends to be misestimated.[21] Accordingly, it is commonly admitted that hemorrhages could be much better defined by a fall in hemoglobin or a need for blood transfusion,[22] with hemorrhages said to be 'major' if the amount of blood needed to be transfused exceeds a certain threshold.[23] This is why the Scottish Confidential Audit of Severe Maternal Morbidity (SCASMM) considered MOH to be defined by a blood loss >2500 ml (definition of its Healthcare Commission) or a transfusion of 5 or more units of blood.[24] The French National Authority for Health (HAS) has concluded that MOH can be defined by any of the following criteria: (i) a blood loss ≥1500 ml, (ii) a hemoglobin loss ≥4 g/dl or (iii) the dispensation of at least 4 units of blood. Therefore, in this study, we operationally defined MOH as cases where 4 or more units of blood were transfused. Although this lack of a uniformly accepted definition could make it more difficult to compare data from different countries, it does not affect the importance or accuracy of MOH monitoring within a single country.

Almost all French women (99.5%) are hospitalized for childbirth,[25,26] and each facility is legally obliged to produce a discharge abstract for each hospital stay. All of these abstracts are included in a national administrative database. Moreover, in France, healthcare expenditure is directly paid to the facilities by the national Social Security Bureau, which relies on these administrative data to calculate expenditure. Therefore, the French administrative database implicitly covers the entire French population, including all age groups. This feature is very rare in other countries. In addition, the quality of national administrative data is regularly monitored.

Here, we aimed to determine whether the national administrative data could be used to estimate the incidence of *post partum* hemorrhage at the national level. To validate the administrative data concerning transfusions, we compared them with a comprehensive database, managed by the national French Blood transfusion Service (FBS). Although the linkage of individual mothers' records between the national administrative database and the FBS database could be an effective way to assess the incidence of hemorrhages, it was not feasible because all of the administrative discharge abstracts are rendered anonymous before being transmitted to the national database. To overcome this difficulty, we relied on the regional database run by the Burgundy Perinatal Network (BPN). Indeed, the perinatal database not only contains data collected from hospitals using the same abstracts as the national database, but also offers the benefit of being linkable with the FBS database. By linking the BPN and FBS data, we aimed to assess the accuracy of discharge abstracts using the FBS database as the 'gold standard'. We then developed a statistical method to correct the inaccurate number of transfusions recorded in the national administrative data

Data sources

The FBS database

In France, the FBS is the single manager of blood collection and distribution. This service acts in compliance with the European directive[27] that sets standards of quality and safety for the collection, testing, processing, storage and distribution of human blood and blood components. It can be compared with the blood services or centers of other countries that are also involved in blood management.[28] This national institution runs a highly reliable and exhaustive database, and it can be safely assumed not only that all blood transfusions are recorded in the FBS registry, but also that only those patients who actually received transfusions are included. Because of this completeness and accuracy, the FBS data could be considered a 'gold standard'. Accordingly, in this study, we searched the Burgundy section of the FBS database to obtain the date of dispensation, the type of blood products (platelets, red blood cells or fresh frozen plasma), the first and last names and date of birth of each woman (pregnant or not) in Burgundy who received a transfusion, as well the name of the hospital where the transfusion was given.

The national administrative database

Inspired by the Medicare DRG system, the gathering of national administrative health data was established in France in 1991(Act of 31 July 1991 on the hospital reform), and

extended in 1997 to all healthcare facilities (Order of 22 July 1996 on the collection and processing of medical data). Initially designed to analyze the hospital activity and to contribute to the elaboration of strategic healthcare plans, it has become an instrument for financial management. Since 2008, each hospital's budget has depended on the medical activity recorded in a specific program that compiles discharge abstracts related to all admissions (Social Security Financing Act for 2008).

The information in these abstracts covers both medical and administrative data, but the latter are rendered anonymous, making it impossible to use the social security number for linkage. This is not usually the case in Medicare data based studies.[29] Diagnoses identified during the hospital stay are coded according to the 10th edition of the International Classification of Diseases. Each facility produces its own anonymous standardized data, which are then compiled at the national level.

Although DRG systems are widely used in the world,[30] the French administrative database offers the attractiveness of being highly exhaustive as it covers the entire population of the country. The fact that these national data are used for the allocation of hospital budgets should encourage improvement in data quality in terms of coherence, accuracy and exhaustiveness. However, the quality and validity of some health-related data, including those regarding blood transfusions, are unknown, and poor accuracy of coding is a frequent limitation of administrative health data.[31] For our purposes, the quality of these data needed to be assessed.

The BPN database

This network was created in 1992 to improve the quality of perinatal care in Burgundy, a French region with 1 800 000 inhabitants and ~18 500 births annually.[32] Since 2001, all of the perinatal care departments in Burgundy have been involved in the BPN, providing data on 100% of the births in the entire region. The BPN regularly collects data (anonymous abstracts) produced by the hospitals for the national administrative database for all pregnant women from 22 weeks of amenorrhea onwards and for all newborns. A small amount of supplementary information is specifically collected by the BPN, such as the gestational age and the delivery date (both of which are now available in the administrative data but were not recorded at the time of the study). Therefore, as the BPN data and the national administrative data come from the same source, assessing the quality of the BPN data can be considered the same as assessing the quality of the national administrative data concerning pregnant women in Burgundy.

For the present study, we selected all births that occurred in Burgundy between 1 January 2006 and 31 December 2007 and were recorded in the BPN database.

Methods

Validation study: assessment of hemorrhage identification at the regional level

In the administrative data (BPN and national administrative database), the notion of 'pregnant woman' is not recorded systematically. Indeed, pregnancy itself does not correspond to a specific diagnosis with a specific code. Of course, all patients included in the BPN database are pregnant, but this limitation has to be taken into account when the national database, which covers the entire French population, is analyzed. Unlike the pregnancy, the delivery is recorded in the administrative data. We thus had to retrieve the discharge abstract related to the hospital stay for childbirth ('birth-related stay') to identify pregnant women [International Classification of Diseases (ICD)-10 codes: Z37, O80 to O84].

Another issue, specific to the administrative data, is that a single record corresponds to a single uninterrupted period of stay in a particular hospital. Therefore, in the case of a patient who was transferred from one hospital to another, each institution will produce its own discharge abstract. For the same reason, if a single pregnant woman was first admitted for hemorrhage then discharged home, and finally re-admitted to the same facility for childbirth, the hemorrhage and the birth will be recorded in two different discharge abstracts. To overcome this difficulty, we linked the abstract retrieved for the birth-related stay with all discharge abstracts for the same patient within the 9 months before childbirth. By doing so, we were able to identify all the relevant pregnancy-related abstracts in order to minimize information loss.

Data anonymization

In accordance with European[33] and French law, both the FBS and BPN data have to be rendered anonymous. As the anonymization method routinely used for the BPN data was developed by our research team (ANONYMAT Software),[34,35] for the specific purpose of this study, we were able to apply the same method to the FBS data. This allowed us to link the BPN and the FBS databases at the individual patient level. ANONYMAT was approved by the French data protection commission for data management in the BPN in general, and specifically for the current study. Its technology is based on the Standard Hash Algorithm,[34,35] which ensures irreversible transformation of independent fields.

Data linkage

To link the BPN with the FBS data, we used the method proposed by Jaro[36] to take into account identification variables, such as maiden names, first names and dates of birth. Technically, there is no difference between using original nominative data or hash-coded data for record linkage, as the correspondence between these data is almost one to one (i.e. low collision rate).

Validation

Figure 1 shows the general flow chart of calculations resulting from comparisons of the BPN and FBS data. False negatives (FN) were defined as BPN abstracts that did not mention a transfusion but for which the FBS database did indicate a transfusion. Conversely, cases where transfusions were recorded in the BPN abstracts, but not in the FBS database, were considered false positives (FP). Once identified, FN and FP were analyzed in detail to determine the causes of the inaccuracy. Regarding the FP, as no information on these patients was provided by the FBS, we had to rely on medical files, whenever possible, in order to understand the reasons for these errors. For FN, we investigated the available FBS data to understand why these transfusions were omitted from the BPN abstracts.

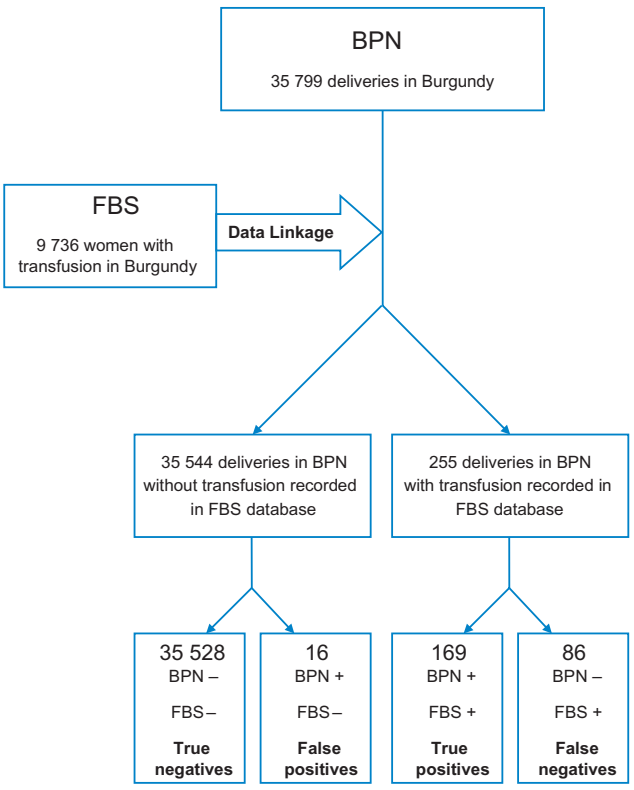

Fig. 1 A flow chart of the linkage of FBS and BPN databases (birth-related stay). BPN+/−, transfusion recorded/not recorded in the BPN database; FBS+/−,: transfusion recorded/not recorded in the FBS database.

When we analyzed the FP and FN, we studied two situations: (i) the *post partum* period alone and (ii) the antepartum and *post partum* periods together. In the former case, we assessed the quality of the administrative data related to a single hospital stay (the birth-related stay), whereas in the latter case, we assessed the potential benefits of linking abstracts related to different stays during the same pregnancy to verify whether this approach could reduce information loss. The sensitivity, specificity and positive and negative predictive values of the BPN abstracts were then estimated separately for each of the two situations.

Furthermore, besides these analyses, for the birth-related stay, we distinguished what can be called 'early' hemorrhages, i.e. those that occurred during the delivery in the delivery room from those that occurred during the rest of the birth-related stay. Early hemorrhages were identified as those cases where the transfusion date (provided in the FBS data) matched the delivery date (provided in the BPN data). By comparing the completeness of recording in the two situations, we aimed to establish whether incidents occurring during childbirth were better recorded than those occurring during the rest of the birth-related stay.

Initially, all of the above-mentioned calculations were made regardless of the quantity of blood delivered. Then, to isolate major hemorrhages, we identified FBS transfusion reports indicating that at least 4 units of red blood cells were transfused, and then estimated the sensitivity, specificity and positive- and negative-predictive values for the corresponding data. By doing so, we aimed to establish whether serious incidents were better recorded than more benign ones.

Modeling study: estimation of the total number of hemorrhages in the national administrative database

This was difficult to estimate because, unlike the BPN database, the national database only records the entire length of a hospital stay, from the date of admission to the date of discharge. The birth date is not recorded. Therefore, we could not accurately retrieve abstracts related to the whole antepartum period as we did in the BPN database. The same applied to early hemorrhages. However, in the national database, each patient is assigned a unique encrypted number. It was thus possible to link all the discharge abstracts for a particular female patient and thus to obtain a rough estimate of the total number of pregnancies during which blood transfusions occurred. We relied on the quality assessment of the BPN administrative data in Burgundy, as

described in the 'validation study' sub-section, to correct the number of transfused women recorded at the national level for the less-than-perfect sensitivity and specificity. In a more simplistic approach, the observed number could merely be corrected for the FN and FP rates as estimated from the Burgundy data. However, such an approach would impose a strong assumption that the frequencies of both types of errors at the national level were identical to those estimated in a single region. In contrast, we expected that the quality of the administrative data would vary systematically with the characteristics of both the women and the health facility, and that these characteristics would, in turn, vary across the regions. Thus, we decided to develop two multivariable regression models that linked the relevant characteristics with the probability of an FN or an FP. In both models, we initially included as the independent variables (i) *a priori* selected clinical severity criteria (mother's age, length of stay and level of the maternity unit), (ii) *a priori* selected geographical and socio-economic factors [geographic area and hospital status (private or public)] and (iii) variables identified by our preliminary validation study (see section Validation; delivery through a Cesarean section, hemostasis, a hemorrhage-related diagnosis and an anemia-related diagnosis). Only the variables found to have statistically significant independent effects were then included in the two separate 'final' multivariable logistic regression models. These models were then used to estimate how the respective probabilities of an FN and an FP depended on the relevant woman's characteristics recorded in the national database.

Each model was estimated using data on a relevant subset of women in the BPN data set, for whom the 'true' transfusion status was known from the FBS. Specifically, the model for FN was estimated using all births with no transfusion recorded in the BPN database. The binary outcome in that model identified the 'FN' subjects. The estimated model was then applied to all births with no transfusion coded in the national database. Specifically, for each of these births, we calculated the estimated probability that a given woman actually had a transfusion, as a function of the values of the aforementioned covariates for that woman. The total number of transfusions missed by the national database (i.e. the expected number of 'FN') was estimated by summing up all of the resulting individual probabilities. Notice that the procedure proposed above takes into account the possibility that the distribution of covariates associated with the probability of a woman having a transfusion may differ between the regional and the national databases.

The variance of the estimated total number of FN depends on the variance and the covariance of the regression coefficients of the logistic model used to estimate the probabilities of FN results. Therefore, the 95% confidence interval (CI) for the total number of FN was estimated based on 500 simulations. In each simulation, the entire vector of logistic regression coefficients was randomly sampled from the multivariate normal distribution, in which both the mean values and the variance–covariance matrix corresponded to the estimates from the original model. For each simulation, the probability that a woman who had a transfusion but was recorded as 'negative' in the national database was re-calculated using the corresponding, randomly sampled, vector of regression coefficients, and the resulting estimate of the total number of FN was obtained as the sum of these probabilities. Finally, the 95% CI for the total number of FN was obtained as the interval between the 2.5th and the 97.5th percentiles of the distribution of the 500 estimates, each corresponding to one simulation.

The same procedure was used to estimate the number of 'FP' in the national database. However, in some situations, the number of FP actually observed at the regional level was considered insufficient for multivariable logistic regression modeling. In such situations, the estimated number at the national level was calculated by simply applying the percentage of FP observed in the regional database to all women identified as 'positive' in the national database.

The total number of transfusions at the national level was estimated by (i) summing the number of transfusions coded in the national database and (ii) adding the estimated number of FN (transfusions missed) and then (iii) subtracting the estimated number of FP, computed as described above. The 95% CI for the estimated total number of transfusions was obtained by combining the estimated variances of the three components of the estimate: (i) the number of transfusions actually recorded in the national database, and the estimated numbers of (ii) FN and (iii) FP. The variance of component (ii) was estimated through simulations, as explained above. The variance of component (iii) was estimated by the conventional method. The SAS macro that implements the above procedure is available on request from the first author.

Results

Results of the validation study

Frequency of transfusions

The FBS database recorded a total of 9736 women (pregnant or not) who received transfusions in 2006 and 2007, in Burgundy.

The BPN database identified 35 799 births during 2006 and 2007, in a total of 35 123 pregnant women. In the corresponding discharge abstracts, we identified a total of 195 transfusions, including 185 that occurred *during* the birth-related hospitalization.

Results of data linkage

The linkage identified 255 transfusions recorded in the FBS database, in a total of 35 799 births, recorded in the BPN database, which represents a 7.12‰ blood transfusion rate. As expected, there were no statistically or clinically significant differences between the number of transfusions recorded in 2006 and 2007 (data not shown).

Table 1 shows the results of analyses conducted on abstracts for birth-related stays recorded in the BPN database, regardless of the quantity of blood delivered. The sensitivity was estimated at 66.3% (60.5–72.1) and the positive predictive value (PPV) at 91.3% (87.3–95.4), with 72.8% (65.7–79.9) of sensitivity and 95.6% of (91.9–99.4) PPV for early hemorrhages. Based on the 'MOH cases' identified in the FBS database, we were able to estimate the frequency of transfusions during which at least 4 units of red blood cells were delivered: the result of these calculations was a 2.3‰ blood transfusion rate. However, as there is no information in the administrative data regarding the quantity of the blood delivered, the sensitivity and the PPV of BPN records could not be estimated regarding MOH.

The results of analyses conducted on linked abstracts related to different stays during the same pregnancy are very similar to the results of analyses conducted on abstracts for birth-related stays.

Review of cases with misclassifications of the transfusion status

For patients identified as 'errors' in the BPN database, all of the medical files available in the respective facilities were reviewed in detail, for a total of 10 FP (77% of all FP) and 78 FN (96% of all FN).

FPs (Table 2) were mostly related to Cesarean section (80%) and many were due to coding errors: either a blood product was mistakenly coded instead of drugs (albumin, coagulation factor XI and complement C1 esterase inhibitor), or no products at all were delivered.

Regarding FN (Table 3), *post partum* hemorrhage was the medical indication for which transfusions seemed to have been recorded with the lowest accuracy. In such cases, most failures to record a transfusion occurred in level-2 maternity units, and mainly for early *post partum* hemorrhage (60.2%), during the two first days after the birth.

Table 1 Sensitivity and positive predictive values of the administrative data to identify transfusions during in-hospital births versus FBS data used as the gold standard in Burgundy.

	Cases identified in the FBS	*Cases identified in the BPN*	*BPN/FBS discordances*		*Sensitivity (%) (95% CI)*	*Specificity (%)*	*Positive predictive value (%)*	*Negative predictive value (%)*	*Rates of transfusion*	
			FPs	*FNs*					*FBS (‰)*	*FBS >4 blood units (‰)*
Deliveries										
Overall	255	185	16	86	66.3 (60.5–72.1)	99.9	91.3 (87.3–95.4)	99.8	7.12	2.3
With early PPH	151	115	5	41	72.8 (65.7–79.9)	99.4	95.6 (91.9–99.4)	95.6	4.22	1.7
Without early PPH	104	70	11	45	56.7 (47.2–66.2)	99.9	84.3 (75.8–92.8)	99.9	2.90	0.6
Pregnancy										
Overall	263	195	13	81	69.2 (63.6–74.8)	99.9	93.3 (89.8–96.8)	99.8	7.35	2.4
With early PPH	155	124	3	34	78.1 (71.6–84.6)	99.7	97.6 (94.9–99.9)	96.3	4.33	1.8
Without early PPH	108	71	10	47	56.5 (47.1–65.8)	99.9	85.9 (77.8–94.0)	99.9	3.02	0.6

Table 2 Distribution of FPs in the BPN database sorted by information provided by patient medical record

	Level 1[a]	*Level 2*[a]	*Level 3*[a]	*Total false positives*
Medical information provided by patient's record				
Anemia	0	0	1	1
Postpartum hemorrhage	1	1	0	2
Hypoproteinemia	0	0	1	1
Hereditary angioedema	0	0	1	1
Factor XI deficiency	0	0	1	1
No indication	2	1	1	4
Total	3	2	5	10
Types of products delivered				
Pharmacy products (drugs)	0	0	3	3
No products issued	2	1	1	7
Total	3	2	5	10
Type of birth				
Cesarean section	3	2	3	8
Vaginal delivery	0	0	2	2
Total	3	2	5	10

1, normal birth at term; 2, birth from 34 weeks of gestation; 3, all births.
[a]Hospital level.

Table 3 Distribution of FNs in the BPN database sorted by information provided by patient medical record

	Level 1[a]	*Level 2*[a]	*Level 3*[a]	*Total FNs*
Indication for transfusion				
Anemia	0	5	1	6
Thrombocytopenia	0	7	4	11
Postpartum hemorrhage	19	31	11	61
Total	19	43	16	78
Date of transfusion				
Antepartum	0	12	5	17
Within 2 days after birth	12	27	8	47
Later	7	4	3	14
Total	19	43	16	78
Type birth				
Cesarean section	8	25	11	44
Vaginal delivery	11	18	5	34
Total	19	43	16	78

1, normal birth at term; 2, birth from 34 weeks of gestation; 3, all births.
[a]Hospital level.

Results of the modeling study

The national administrative database identified 1 629 537 deliveries between 2006 and 2007, with a total of 6932 (0.43%) blood transfusions.

Table 4 shows variables that were significantly associated with the probability of an FN in the BPN database, based on the multivariable logistic regression model (AUC = 0.907). These variables were mainly related to the patient, and included Cesarean section, hemorrhage, anemia, hemostasis procedures and a longer length of stay. These variables have been included in the final predictive models, employed to correct the number of transfusions recorded in the national administrative database. In contrast, mother's age, geographical area, hospital status and the level of the maternity unit had no statistically significant associations with the probability of FN and were thus not included in the final predictive model used to correct the national data. The total number of FN at the national level, estimated according to the predictive model, was 4609 (95% CI: 3990–5996). By adding this number to the 6932 observed and subtracting the estimated 600 FP, we calculated that the correct number

Table 4 Results of the multivariable logistic regression model predicting 'false-negative' results from the BPR database

Characteristics	*Adjusted OR*	P-*value*
Intercept	−7.9493	<0.0001
Cesarean Section		
Yes	1.0755	0.0364
Hemorrhage		
Yes	3.1019	<0.0001
Anemia		
Yes	1.7231	<0.0001
Length of stay		
0–3 days	−0.046	0.9554
6–14 days	1.6632	0.0008
>14 days	3.0375	0.0001
Hemostasis procedure		
Yes	6.0343	<0.0001
Hemorrhage × length of stay		
Yes		
0–3 days	1.954	0.0353
6–14 days	−0.5308	0.3804
>14 days	−1.9832	0.0738
Cesarean section × length of stay		
Yes		
0–3 days	2.3337	0.0216
6–14 days	−0.4214	0.5235
>14 days	−0.7049	0.4751

of pregnant women who received transfusions in 2006–07 in France was 10 941 (95% CI: 9899–11 983), i.e. 671.4 per 100 000 deliveries.

Discussion

In this article, we aimed to assess the feasibility of using the national administrative database, which contains data collected primarily for hospital budget allocation, to monitor obstetric hemorrhages. To assess the quality of the administrative data for blood transfusions, in principle, we could rely on the 'gold standard' of the highly reliable national FBS database of blood delivery. However, due to the anonymization procedures applied to the national administrative data, it was not possible to link them to the FBS data. Nevertheless, in the region of Burgundy, we were able to link the FBS data to the administrative data maintained by a regional perinatal network (BPN). The results of this local assessment were incorporated into two multivariable regression models to estimate the number of pregnant women who received a transfusion according to the national administrative database. It is important to note that during the assessment of the quality of transfusion recording at the regional level we relied on an accurate and validated record of pregnant women who received a transfusion, so that neither extrapolation nor statistical estimation was used at this stage. The information that was obtained can thus be considered a reference.

Our results showed that whether all of the relevant pregnancy-related abstracts (69.2% sensitivity) or only the birth-related abstracts (66.3% sensitivity) were taken into account, both the completeness and the accuracy of transfusion recording did not change significantly. In both situations, the rate of transfusions was seriously under-estimated. Furthermore, our analyses indicated the lack of the (expected) relationship between the completeness of transfusion recording and the place where the hemorrhage occurred. Indeed, the sensitivity did not differ significantly when hemorrhages occurring in the delivery room were separated from those occurring during the rest of the birth-related stay. As for transfusions that required at least 4 units of red blood cells, there was no information in the administrative data regarding the quantity of blood delivered. Thus, in order to make the administrative data potentially useful in MOH monitoring, this data limitation has to be overcome.

The overall poor sensitivity of the BPN data confirms the findings of another validation study of French hospital discharge database for severe maternal morbidity,[31] in which the coding validity of diagnosis was described as 'poor'. However, after assessing the quality of the BPN administrative data, our main goal was to use the resulting estimates of FP and FN rates to correct the number of transfused pregnant women recorded at the national level. To this end, we employed multivariable regression to link the quality of the administrative data with the characteristics of the patient and hospital. The results indicated that factors with a significant impact on data quality were mainly related to the patient ('Cesarean section',' hemorrhage', 'anemia', 'hemostasis procedures') and a longer length of stay. In contrast, the age of the patient and criteria concerning the hospital ('geographical area', 'hospital status: private or public', 'facility level') were not significantly associated with data quality. These results suggest that possible variations in the organization of the different facilities across the regions are unlikely to affect the accuracy of the results of our extrapolation from the regional to the national level.

Based on the multivariable models, we estimated that the corrected transfusion rate was 671.4 per 100 000 deliveries, i.e. 0.671%. We compared our estimate with the results of a cluster-randomized trial that involved 106 French maternity units.[37] Indeed, although the estimate provided by modeling of population-based data may not be as reliable as the one provided by a randomized trial, the aforementioned trial reported blood transfusion rates close to our estimate: 0.44% of deliveries in the intervention group and 0.41% deliveries in the control group On the other hand, while the accuracy of data collection in the trial was better than in administrative databases used in our study, the representativeness of the study sample was necessarily lower than in a population-based study. Our result falls within the range of estimated transfusion rates (between 0.34 and 0.90% of deliveries) reported by other French authors,[38,39] and is also very similar to the transfusion rate found in Canada (0.647% of deliveries).[40,41]

Finally, we aimed to demonstrate that, despite the considerable inaccuracy of coding in the national administrative data, routinely collected data can provide accurate estimates if corrected by multivariable regression models that account for the variables significantly associated with the probability of coding errors. Thus, combining the administrative database with careful modeling can yield accurate estimates while having the advantage of describing the general population. In fact, the major rationale for our study was to develop and implement the methods that can correct for the limited sensitivity of the administrative records of transfusions, in which transfusions are substantially under-reported. Indeed, the results presented in Table 1 and discussed in the paper show that our method increased the estimated total number of transfusions from 6932 to 10 941, i.e. by ∼42% of the estimated (corrected) total number.

This 42% increase is quite consistent with the overall sensitivity of the administrative database of ~60% (which implies that ~40% of 'true' cases will be missed).

Conclusion

The administrative data on blood transfusion can be used to assess the frequency of obstetric hemorrhages in any country where a large administrative database is available. In France, some limitations still remain concerning major hemorrhages, as there is no information in the administrative database regarding the quantity of blood delivered. Providing such information would increase the effectiveness of our method.

One way would be to add a new field in the administrative database to record the quantity of blood delivered. Indeed, the information system used[42] to gather and store such data has been designed and developed so that new information is regularly added. Moreover, if errors do occur in the reporting of this information, the model presented in this paper is designed to take such errors into account. However, if quantity data were included, the software publishers would also be required to amend and extend their data capture software accordingly, and all the health professionals would be asked to record this information whenever needed. Such an increase in the workload would only be permitted by the authorities if it is justified by a positive economic impact, i.e. if recording the quantity of blood delivered would be useful to adjust the budgets of individual hospitals.

Alternatively, the linkage between the administrative database and the blood transfusion service database is technically feasible, for example thanks to the National Health Insurance Information System,[42] which allows different national health databases to be linked together. Moreover, this kind of linkage has been allowed by the French data protection commission.[43,44] Therefore, it would neither be a technical issue, nor a matter of authorization, but rather a problem of political will. As long as this linkage is acceptable to healthcare facilities, this alternative approach could be of great interest to assess the frequency of transfusions—and not only obstetric transfusions—at the national level.

Acknowledgements

The authors thank Ms Rebecca Burne, MSc from McGill University for her assistance in the revisions of the manuscript. Michal Abrahamowicz is a James McGill Professor of Biostatistics at McGill University.

References

1 Ronsmans C, Graham WJ. Maternal mortality: who, when, where, and why. *Lancet* 2006;**368(9542)**:1189–200.

2 Stones W, Lim W, Al-Azzawi F *et al*. An investigation of maternal morbidity with identification of life-threatening 'near miss' episodes. *Health Trends* 1991;**23(1)**:13–5.

3 Mantel GD, Buchmann E, Rees H *et al*. Severe acute maternal morbidity: a pilot study of a definition for a near-miss. *Br J Obstet Gynaecol* 1998;**105(9)**:985–90.

4 Hill K, Thomas K, AbouZahr C *et al*. Estimates of maternal mortality worldwide between 1990 and 2005: an assessment of available data. *Lancet* 2007;**370(9595)**:1311–9.

5 Lewis G (ed.). The confidential enquiry into Maternal and Child Health (CEMACH). The Seventh Report on Confidential Enquiries into Maternity Deaths in the United Kingdom: 2003–2005. London: CEMACH, 2007.

6 Schutte JM, De Boer K, Briët JW. Maternal mortality in the Netherlands: the trip of the iceberg. *Ned Tijdschr Obstet Gynaecol* 2005;**118**:89–91.

7 Atrash HK. Not all prenatal care is appropriate care. *J Am Med Womens Assoc* 1995;**50(5)**:144–6.

8 Betran AP, Wojdyla D, Posner SF *et al*. National estimates for maternal mortality: an analysis based on the WHO systematic review of maternal mortality and morbidity. *BMC Public Health* 2005;**5**:131.

9 Khan KS, Wojdyla D, Say L *et al*. WHO analysis of causes of maternal death: a systematic review. *Lancet* 2006;**367(9516)**:1066–74.

10 Gissler M, Deneux-Tharaux C, Alexander S *et al*. Pregnancy-related deaths in four regions of Europe and the United States in 1999–2000: characterisation of unreported deaths. *Eur J Obstet Gynecol Reprod Biol* 2007;**133(2)**:179–85.

11 Schutte JM, Steegers EA, Schuitemaker NW *et al*. Rise in maternal mortality in the Netherlands. *BJOG* 2010;**117(4)**:399–406.

12 Schutte JM, de Jonge L, Schuitemaker NW *et al*. Indirect maternal mortality increases in the Netherlands. *Acta Obstet Gynecol Scand* 2010;**89(6)**:762–8.

13 Van Roosmalen J, Zwart J. Severe acute maternal morbidity next term in high-income countries. *Best Pract Res Clin Obstet Gynaecol* 2009;**23**:297–304.

14 Callaghan WM, Mackay AP, Berg CJ. Identification of severe maternal morbidity during delivery hospitalizations, United States, 1991–2003. *Am J Obstet Gynecol* 2008;**199(2)**:133 e1–8.

15 Pallasmaa N, Ekblad U, Gissler M. Severe maternal morbidity and the mode of delivery. *Acta Obstet Gynecol Scand* 2008;**87**:662–8.

16 Wen SW, Huang L, Liston R *et al*. Severe maternal morbidity in Canada, 1991–2001. *CMAJ* 2005;**173(7)**:759–64.

17 Chang JJ, Stamilio DM, Macones GA. Effect of hospital volume on maternal outcomes in women with prior cesarean delivery undergoing trial of labor. *Am J Epidemiol* 2008;**167(6)**:711–8.

18 Klemmensen AK, Olsen SF, Osterdal ML *et al*. Validity of preeclampsia-related diagnoses recorded in a national hospital registry and in a postpartum interview of the women. *Am J Epidemiol* 2007;**166(2)**:117–24.

19 Brace V, Penney G, Hall M. Quantifying severe maternal morbidity: a Scottish population study. *BJOG* 2004;**111(5)**:481–4.

20 Healthcare Commission Towards better births: a review of maternity services in England. http://webarchive.nationalarchives.gov.uk/20110504133205/http:/www.cqc.org.uk/_db/_documents/Towards_better_births_200807221338.pdf (21 March 2012, date last accessed).

21 Schorn MN. Measurement of blood loss: review of the literature. *J Midwifery Womens Health* 2010;**55(1)**:20–7.

22 Waterstone M, Bewley S, Wolfe C. Incidence and predictors of severe obstetric morbidity: case-control study. *BMJ* 2001; **322(7294)**:1089–93; discussion 93–4.

23 van Roosmalen J, Zwart J. Severe acute maternal morbidity in high-income countries. *Best Pract Res Clin Obstet Gynaecol* 2009; **23(3)**:297–304.

24 Rains F, Archibald J, Deyhle D. Scottish Confidential Audit of Severe Maternal Morbidity: 7th Annual Report. *Reweaving the World The Emergence of Ecofeminism* 2011:73–86.

25 Barbier ML, Sanchez P, Cortesi C. Naissances inopinées à domicile en Ile de France. *Rev SAMU* 1990;**2**:69–74.

26 Duroy E, Manzona C, Adamia C *et al*. Evaluation retrospective des accouchements inopinés pris en charge par le SMUR. Urgences 2009–3eme congrès de la Société Française de Médecine d'Urgence; 2009. *J Européen des Urgences* 2009;**22**:A91.

27 Directive 2002/98/CE of the European Parliament and the Council of 27 January 2003 setting standards of quality and safety for the collection, testing, processing, storage and distribution of human blood and blood components and amending Directive 2001/83/EC, 2003.

28 Rock G, Akerblom O, Berseus O *et al*. The supply of blood products in 10 different systems or countries. *Transfus Sci* 2000; **22(3)**:171–82.

29 Virnig B, Durham SB, Folsom AR *et al*. Linking the Iowa Women's Health Study cohort to Medicare data: linkage results and application to hip fracture. *Am J Epidemiol* 2010;**172(3)**:327–33.

30 Quantin C, Benzenine E, Abbas R *et al*. *The Usefulness of Collecting Data from Discharge Abstracts to Estimate Cancer Incidence*. Data Collection and Storage Chapters (chapter 1), Editor Nova Publishers, 2012, 1–26.

31 Chantry AA, Deneux-Tharaux C, Cans C *et al*. Hospital discharge data can be used for monitoring procedures and intensive care related to severe maternal morbidity. *J Clin Epidemiol* 2011; **64(9)**:1014–22.

32 Quantin C, Gouyon B, Avillach P *et al*. Using discharge abstracts to evaluate a regional perinatal network: assessment of the linkage procedure of anonymous data. *Int J Telemed Appl* 2009:ID 181847.

33 Directive 95/46/CE of the European parliament and of the council of 24 october 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995.

34 Quantin C, Bouzelat H, Allaert FA *et al*. Automatic record hash coding and linkage for epidemiological follow-up data confidentiality. *Methods Inf Med* 1998;**37(3)**:271–7.

35 Quantin C, Bouzelat H, Allaert FA *et al*. How to ensure data security of an epidemiological follow-up: quality assessment of an anonymous record linkage procedure. *Int J Med Inform* 1998;**49(1)**:117–22.

36 Jaro MA. Probabilistic linkage of large public health data files. *Stat Med* 1995;**14(5–7)**:491–8.

37 Deneux-Tharaux C, Dupont C, Colin C *et al*. Multifaceted intervention to decrease the rate of severe postpartum haemorrhage: the PITHAGORE6 cluster-randomised controlled trial. *BJOG* 2010;**117(10)**:1278–87.

38 Francois A, Courtois F. Management of blood products in the event of postpartum hemorrhage. *J Gynecol Obstet Biol Reprod (Paris)* 2004;**33(8 Suppl.)**:4S120–4S9.

39 Reyal F, Sibony O, Oury JF *et al*. Criteria for transfusion in severe postpartum hemorrhage: analysis of practice and risk factors. *Eur J Obstet Gynecol Reprod Biol* 2004;**112(1)**:61–4.

40 Ickx BE. Fluid and blood transfusion management in obstetrics. *Eur J Anaesthesiol* 2010;**27(12)**:1031–5.

41 Liu S, Joseph KS, Bartholomew S *et al*. Temporal trends and regional variations in severe maternal morbidity in Canada, 2003 to 2007. *J Obstet Gynaecol Can* 2010;**32(9)**:847–55.

42 Goldberg M, Jougla E, Fassa M *et al*. The French health information system. *J LAOS* 2012;**28(1)**:31–41.

43 Trouessin G, Allaert FA. FOIN: a nominative information occultation function. *Stud Health Technol Inform* 1997;**43 (Pt A)**: 196–200.

44 Vuillet-Tavernier S. Réflexion autour de l'anonymat dans le traitement des données de santé. *Médecine et Droit* 2000;**40**:1–4.