Multi method approach to the assessment of data quality in the Finnish Medical Birth Registry

J Teperi

Abstract

Objective—To assess comprehensively the validity of the data in the Finnish Medical Birth Registry (MBR) by the combined use of several controls and internal analysis of the data.

Design—The MBR data were individually linked to a medical record sample (n=775) and to all perinatal death certificates in 1987. The data were also compared with annual hospital statistics. The distributions of birth weights and gestational ages were examined.

Subjects—All stillborn and liveborn babies registered in the MBR in 1987 (n=59 370).

Setting—The nationwide MBR data were compared with medical records from one third of the Finnish hospitals, with statistics for all hospitals, and with nationwide cause of death registry data.

Measurements and main results—With regard to most variables, the data quality was good or satisfactory (agreement with medical records 95% or more). Allowing for minor deviations in variables with continuous scales improved the agreement rates further. Explanations could be deduced for items with poor agreement values. For most variables, the amount of missing data was less than 1%. With the exception of caesarean sections, medical procedures were registered in only 30 to 72% of the cases, and the proportion varied strongly between the hospitals. Common diagnoses (32 to 86%) and primary causes of death (59 to 78%) were also poorly recorded.

Conclusions—Combined use of several control materials and internal analyses was successful in investigating the whole data content. The data in the MBR were generally valid but diagnoses and most data on medical procedures were not of sufficiently good quality.

J Epidemiol Community Health 1993; 47: 242–247

Department of Public Health, Haartmaninkatu 3, 00290 Helsinki, Finland
J Teperi

Correspondence to:
Dr J Teperi

Accepted for publication October 1992

The key issue in making large medical information systems cost effective is the utility of the gathered data. In many registries on perinatal health, a high rate of non-recording among selected subpopulations has often caused a fundamental data quality problem. In data collection systems in which more than 20% of neonates who die are not registered, for example,[1] it is obvious that the utility of the data is seriously compromised.

Another problem reported repeatedly is the validity of the recorded information in perinatal databases. For example, data on one of the most important correlates of newborns' health, the length of gestation, are almost useless when 84% of all births are recorded as taking place at either 36 or 40 weeks.[2] Research on specific diseases responsible for stillbirths is hampered if fewer than one quarter of the diagnoses on fetal death reports are relevant.[3]

The Finnish Medical Birth Registry (MBR) was established in 1987 by the National Board of Health. In a data quality study before this report, it was shown that 1·5% of newborns were not recorded in the MBR in 1987.[4] Perinatal deaths were over-represented among those neonates who were not registered.

This study was conducted to investigate the validity of the data recorded in the new information system on perinatal care and health. It was decided to combine several control materials plus internal analysis of the MBR data to obtain a comprehensive view of the quality of the data. The study provides a basis for improving data quality, and it also serves as a baseline reference for future assessments of validity. In the discussion section, special emphasis is given to methodological issues and implications that can be generalised to other medical registries.

Methods

Since 1 January 1987, the National Board of Health (NBH) has run the nationwide MBR in Finland. Hospitals are responsible for reporting all stillborn and liveborn babies on an MBR data sheet, which is to be filled in not later than seven days after birth. A ward clerk or midwife usually completes most of the details, and a doctor records any diagnostic data.[5]

In 1987, most hospitals transferred their data monthly to the NBH on the original data sheets. At the NBH the data were computerised and then sent to the State Computing Centre (SCC) on disks. Twenty hospitals sent their data every six months on magnetic tapes directly to the SCC. After a routine data validation procedure, consisting of locating extreme or otherwise implausible values, internal inconsistencies, and missing data in the most important variables, an error listing of the annual data was sent to all hospitals. They were expected to return the corrections within a few months. The data content in the MBR, and in essential cases the structure of the questions on the data sheet are presented in the Results section. The total number of newborns recorded in the MBR in 1987 was 59 370.

Table I   Numbers of hospitals and neonates in the Medical Birth Registry in 1987, and in the medical record sample according to the type of hospital

	Medical Birth Registry			Sample		
	Hospitals	Newborns		Hospitals	Newborns	
Type of hospital	no	no	(%)	no	no	(%)
University teaching hospitals	5	17 200	(29)	2	292	(38)
County hospitals	16	21 652	(36)	5	272	(35)
Local hospitals	27	19 900	(33)	9	205	(26)
Health centres	5	607	(1)	1	6	(1)
Births at home		11	—	—	—	—
Total	53	59 370	(100)	17	775	(100)

Three controls were used to assess the validity of the MBR data. The main validation was based on a sample of medical records, and it is called the 'medical record data' in this paper. Twelve days were randomly selected from the second quarter of 1987. The sample consisted of all the babies born on those days in one third of all the hospitals with maternity wards. The hospitals were selected as a stratified sample to ensure that all types of hospital were represented (see table I). After obtaining the required authorisation, the data corresponding to the MBR content were abstracted from 775 medical records from 17 hospitals by one trained research assistant who was also a qualified nurse.

The corresponding data were selected from the MBR by using the personal identification number of the mother as the linkage key variable. Out of the total of 775 abstracted medical records, MBR data could be found for 760 neonates. The number of cases with no data in the MBR did not differ significantly from the average non-recording rate (1·5%) of the registry. There were six babies, who according to the MBR had been born in the sample hospitals during the selected days, but for whom no medical record data were found. These cases did not differ systematically from the rest of the sample. Probably these neonates' medical records were unintentionally omitted by the hospitals' staff.

Most of the information needed to complete the validation was readily found in the medical records. Most variables in the resulting medical record data were complete; in nine variables there were one to five cases with no data available. The variables with more missing data were: 'where discharged to' (12 cases with no information), 'number of antenatal visits' (n=162), 'date of the first antenatal visit' (n=170), and 'the Apgar score at five minutes' (n=171).

To assess the MBR data on procedures, a data set originally collected for an international comparative study[6] was used. In Finland, a questionnaire mailed to heads of all obstetric units was used to collect data on the application of various procedures. All hospitals responded, but not all the hospitals could report all the procedures that were listed on the questionnaire. The data produced were mainly based on maternity ward diaries, which are considered to be reliable. This source is called 'the maternity ward data' here.

The Central Statistical Office maintains a national cause of death registry, based on death certificates. We received the data on death certificates for all stillbirths and early neonatal deaths in 1987. This material is called 'the death certificate data' here, and it was used to evaluate the accuracy of some of the diagnostic data in the MBR.

Results

INTERNAL ANALYSIS OF THE MBR DATA

In the total MBR data (n=59 370), data were missing for about half of the variables in only 1/1000 cases, or fewer. The rate was more than 1% in eight variables as follows: 'where discharged to', 'age at discharge', 'number of antenatal visits', 'marital status', 'date of the first antenatal visit' (less than 1·5%), 'maternal smoking during pregnancy' (2%), 'duration of pregnancy estimated by ultrasound (yes/no)' (8%), and 'Apgar score at five minutes' (16%).

The distributions of birth weight and length of gestation were analysed because of the importance of these variables. Birth weight data were missing for only 23 newborns (0·04%). Forty five of the 48 newborns with extreme birth weights <500 g and >5500 g) had consistent data, including diagnoses, supporting the accuracy of the birth weight data.

The overall distribution of the birth weight data showed a regular, narrowish bell shaped configuration (mean (SD) 3549·7 g (583)), with an anticipated slight left skew (skewness −0·89) caused by preterm births. The gestation week

Table II   Agreement between the Medical Birth Registry and medical record data (total n=760) (data on procedures and diagnoses excluded)

Good agreement	Non-agreement Rate (%)	*
Ordinal sign (in plural births, n=10)	0·0	
One minute Apgar score	0·0	
No of fetuses (in plural births, n=10)	0·0	
First/second fetus born alive/dead (in plural births)	0·0	
Time of death (in early neonatal deaths, n=4)	0·0	
Liveborn/stillborn	0·1	
Hospital code	0·3	
Sex	0·3	
No of previous stillbirths	0·4	
+/−1 stillbirth		0·0
Birth weight	0·8	
+/−100 g		0·1
No of previous livebirths	0·9	
+/−1 livebirth		0·8
No of previous extrauterine pregnancies	1·2	
+/−1 extrauterine pregnancy		0·1
Five minute Apgar score	1·4	
No of previous pregnancies	1·6	
+/−1 pregnancy		0·5
No of previous induced abortions	1·9	
+/−1 abortion		0·0
Satisfactory agreement:		
Marital status	2·5	
Place of residence	2·6	
Time of birth	2·8	
+/−30 minutes		0·7
No of previous miscarriages	2·9	
+/−1 miscarriage		0·7
Date of the first antenatal visit	3·3	
+/−30 days		2·1
Cohabiting	3·8	
Maternal smoking during pregnancy	4·8	
Goods or satisfactory agreement, when relaxed conditions applied:		
Birth length	25·9	
+/−1 cm		0·7
No of antenatal visits	12·2	
+/−2 visits		2·7
Age at discharge	57·3	
+/−1 day		3·3
Duration of labour†	38·7	
+/−30 minutes		3·9
Unsatisfactory agreement:		
Duration of pregnancy	17·4	
+/−6 days		6·1
Where discharged to	13·7	
Estimate on duration of pregnancy by ultrasound, yes/no	29·7	
Time of death in stillbirths (n=3)	33·3	

* Number of cases with discrepant data in MBR and medical record data, divided by number of cases with data available in both data sources, and multiplied by 100.
† Only for vaginal births (n=644)

specific weight distributions between 38 and 41 weeks of gestation were almost symmetrical (skewness ranging from 0·07 to 0·15). A round digit preference was observed as an over-representation of even 100 g (74% more than the expected number) or even 50 g (31%).

The data on gestational age were missing in 139 records (0·2%). These were mostly normal and healthy babies, but a non-significant over representation of newborns with a birth weight less than 1500 g was observed. There were 38 (0·06%) neonates born at term with a one position move error in data entry. In addition, 87% of the 79 extreme gestations (5–18 weeks or at least 44 weeks) were obvious errors.

It has been repeatedly reported elsewhere, that the gestation week specific weight data in the preterm period show a bimodal distribution.[7] This has been attributed to misclassification of some term births as preterm ones. This could not be demonstrated in our data.

The total MBR material was also used to evaluate the representativeness of the MBR sample, which was linked to the medical record data. In variables where more than 1% of values were missing in the whole registry (n=8), the proportion of blanks in the sample was 1·2 to 2·8 times higher. In the 'duration of gestation' data, the 1·3% MBR sample included 55% of all the one position data entry errors, resulting in a clear over estimate of error frequency.

COMPARISON WITH MEDICAL RECORD DATA
The agreement between the MBR and medical record data, excluding diagnoses and procedures, is summarised in table II. For most variables, agreement was good, or satisfactory after allowing for a slight discrepancy in variables with continuous scales.

There was usually a plausible explanation for the variables with poor quality data. The reasons were either technical or related to inadequate instructions given to the MBR data recorders. Almost all the errors in length at birth were caused by rounding half centimetres down to the next smaller integer. The same sort of error occurred with 'age at discharge'. In 'where discharged to (newborn)', the alternatives given were not mutually exclusive, and there was no clear information indicating when the question should be answered. The standard method in hospitals is to include all the three stages of labour when reporting the duration of labour. Only the durations of first and second stages of labour were to be recorded in the MBR, however, which resulted in confusion.

In 6% of the cases, the estimate of duration of pregnancy was wrong by more than six days. About half of these mistakes were caused by an error in data entry, resulting in a one position move in the data record. Up to two thirds of the remaining discrepant cases were the result of recording differing estimates based on alternative methods. No method of estimating the length of pregnancy was given preference in the instructions.

Procedures used during pregnancy (maximum number four), during labour (maximum two), and to the newborn (maximum seven) were recorded in the MBR using codes given in a list on the other side of the MBR data sheet. In procedures used

during labour, there were also two additional check-box items, induction of labour, and caesarean section.

Because of the small size of the medical record sample, only the most common procedures could be analysed (table III). Usually, a half to three quarters of the procedures reported in the medical records were also found in the MBR. In general, the procedures used early in the pregnancy, and those mainly undertaken by midwives rather than doctors were entered less completely in the MBR. Data on caesarean sections were almost complete probably because of the nature of this major intervention, and the specific check-box question used to record it.

Table III Completeness of reporting (%) frequent procedures* in the Medical Birth Registry (MBR) compared with the medical record data (MRD)

Procedure	No in MRD	% in MBR (95% CI)	
Caesarean section	116	97	(91,99)
Epidural anaesthesia during vaginal birth	66	77	(65,87)
Paracervical blockade	115	72	(64,80)
Vacuum extraction	46	70	(54,82)
Induction of labour	123	69	(61,77)
Phototherapy	69	61	(48,72)
Amniocentesis	53	57	(42,70)
Electronic monitoring	644	51	(47,55)
Amnioscopy	125	50	(42,59)
Episiotomy	433	48	(43,52)
Suturation of laceration	53	30	(18,44)

* Procedures with more than 40 cases in the medical record data

ICD-9 codes were used to record diagnoses in the MBR. In most cases, 60–70% of the frequent diagnoses in the medical record data were also found in the MBR (table IV). The worst level of completeness was observed among those diagnoses made early in pregnancy which had no potential effects on the care during the hospitalisation period for labour, or those with somewhat ambiguous diagnostic criteria. The application of three digit ICD categories made the agreement somewhat better.

Consideration of the agreement values only does not give the whole picture of the utility of various variables. When the problems are encountered predominantly in a certain category (or categories) of a variable, or when variables with some invalid data are used together, the problems may multiply. For example, the agreement of both 'marital status' and 'maternal smoking during pregnancy' was more than 95%. However, 35% of the unmarried smokers could not be identified in the MBR data. This was mainly explained by the poor sensitivity in detecting unmarried women and smokers in each appropriate variable.

Errors in various variables were evenly distributed among the cases, but the missing values tended to cluster in the same records. When the variables were pooled together, the rate of disagreement was not associated with the age, marital status, or socioeconomic status (a three category variable based on the medical record data on occupation) of the mother. There was a significant association, however, between missing data and mother's age (from the youngest to the oldest tertile 9%, 6%, and 3% had at least one missing value; p=0·018, test of a linear trend in proportions). The difference was not attributable to any single variable. A similar effect could not be

Table IV *Completeness of reporting frequent diagnosies* in the Medical Birth Registry (MBR) compared with the medical record data (MRD)*

Diagnosis (ICD-9 code)	No in MRD	% in MBR with same code	% in MBR with same three digit category
Fracture of clavicle (7672A)	29	86	86
Breech presentation (6522A)	24	67	88
Abnormal glucose tolerance in pregnancy (6488A)	24	67	75
Unspecified jaundice (7746X)	57	67	72
Mild pre-eclampsia (6424A)	22	64	73
Preterm infant (excluding extreme immaturity) (7561B)	31	61	61
Threatened labour (6440A)	71	32	34
Genitourinary infection in pregnancy (6466A)	47	25	25

* Diagnoses with more than 20 cases in medical record data included

shown in the medical record data. The staff who recorded data in the MBR had, thus, failed to transfer all the data already filed in the medical records of young mothers.

COMPARISON WITH MATERNITY WARD DATA
In hospitals in which maternity ward data were available, the average proportion of caesarean sections found in the MBR was 97%. The same figure for vacuum extractions and episiotomies were 75% and 67%, respectively. The figure shows the wide variation in the completeness of reporting between hospitals, which reduces substantially the utility of the data on medical procedures.
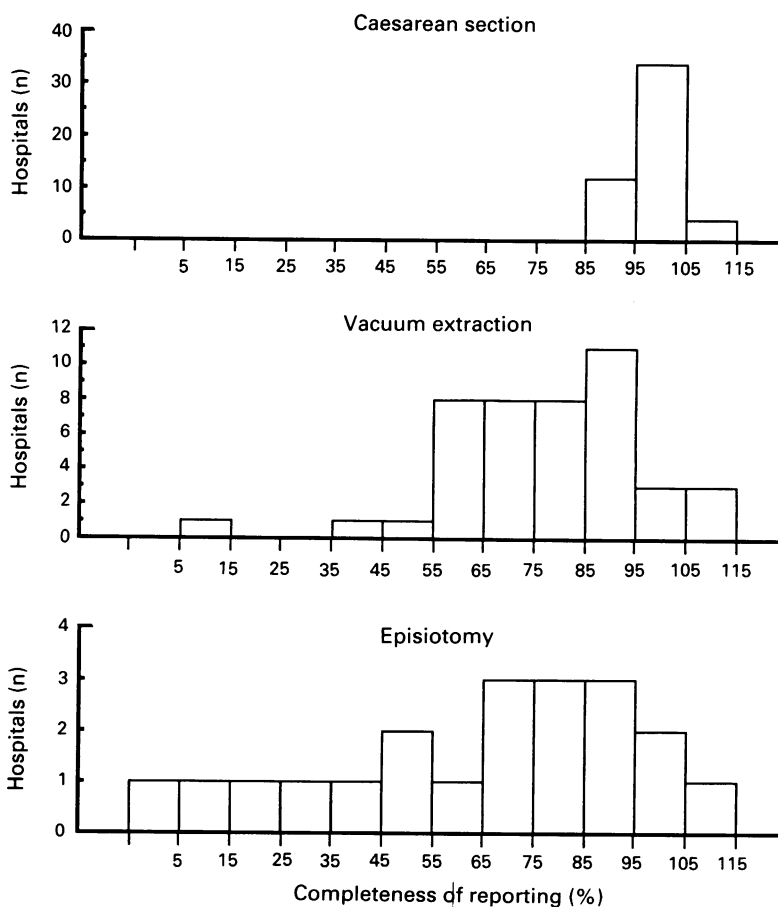

Figure *Distribution of hospitals according to the completeness of reporting of some procedures. Completeness (%)=the number of the procedure in MBR divided by the number in maternity ward data, and multiplied by 100. The data for caesarean section were available from 50 hospitals, for vacuum extraction from 44 hospitals, and for episiotomy from 20 hospitals.*

COMPARISON WITH DEATH CERTIFICATES
The data on the primary causes of death recorded in the death certificate data were compared with diagnoses recorded in the MBR. For stillborn babies (n=256), 59% of the primary causes of death were found in an identical form in the MBR, 18% showed a very close diagnosis, 20% were discrepant, and for 3% no diagnoses were found that could be related to death. For those newborns who died during the first week of life (n=186), the corresponding figures were 78%, 4%, 17%, and 2%, respectively. The vast majority of the discrepant diagnoses, especially among the early neonatal deaths, were the result of more specific diagnosis in the death certificate that was apparently obtained after the necropsy.

Discussion
Many methods have been used to study the quality of data in medical registries. Abstracting from existing documents, medical records or death certificates as typical examples, is probably the most common method used. There are numerous studies on the reliability of the abstracting procedure,[8-10] or on the validity of these reference sources.[11-15] It seems to be widely accepted that abstracting can be done correctly enough after the careful training of abstractors to follow clearly defined criteria, preferably combined with reliability assessment. The major concern is the validity of medical records. Since a failure to record some of the diagnoses and procedures seems to be typical in medical records, it has been assumed traditionally that this method systematically underestimates error frequency in these variables.[16]

Survey methods have also been adapted to obtain the standard material. It is obvious, however, that the potential bias caused by selected non-response seriously reduces the feasibility of such an assessment.[17]

Standard materials can also be obtained without costly and time consuming primary data collection. When data in more than one computerised registry relate to the same person or event, it is relatively economical to link them and analyse the data concordance individually.[18] The critical issue is the existence of unambiguous identifiers in all data sources. The value of this approach is compromised when there are potential selection mechanisms related to a substantial non-linkage rate.

An economical way is to use existing aggregate data such as hospital or administrative area statistics (for example,[19]). The usefulness of this method is restricted by the limited scope of existing statistics, and the impossibility of analysing inconsistencies individually.

Some insight into data quality can also be obtained by careful examination of the distribution of observations, with no external controls. This approach has been successfully used where an expected distribution is known a priori, like birthweight of newborns,[7] or where there are very large differences in distributions between subpopulations of the registry, such as hospital specific diagnose rates.[20]

All the approaches mentioned above, except the use of survey techniques, have been used to study

data quality in the Finnish MBR. (A comprehensive study on the completeness of recording, using automated record linkage to population registration, has been reported elsewhere.[4]) The gains of combining several techniques are demonstrated in this study. Abstracting medical records made it possible to cover all the variables included in the MBR. Because of limits set by an affordable sized sample, however, rare phenomena and data quality variation between all hospitals could be studied only through the additional materials. This was done with relatively small additional resource input. To some extent, the methods could also be used to validate each other.

According to earlier reports, the validity of the data in perinatal registries varies considerably according to the variables examined. The basic demographic characteristics,[16 20 22] and birth weight[7 17 20 23] are usually registered reliably. A wide variation in data quality seems to be typical, however, for information on previous pregnancies,[17 24] the duration of the index pregnancy,[7 17 20] and medical procedures during the pregnancy and labour.[16 19 20 24] Data on prenatal care,[17 21 24] and especially any diagnostic data,[3 16 25 27] tend to be of inferior quality in almost all registries for which data quality studies have been published.

In our study the rate for missing data was very low, and the error frequency in many variables was considerably less than in most reports from other registries. Data on most variables was sufficiently accurate and complete for anticipated research activities. Data on key parameters such as length of gestation and, especially, birth weight, were of excellent quality. In the few variables for which validity was poor, an obvious reason for the problems could be spotted. These problems were taken into account when the MBR data content and the instructions to the data recorders were modified in 1990.

Not surprisingly, the diagnostic data were found to be of poor quality. Although there was variation in the completeness of data found according to the type of diagnosis, even the fundamental data on the primary cause of death were lacking in roughly 20% of perinatal deaths.

Another weak point in the MBR was the recording of data on medical procedures. With the single exception of caesarean section, it seems that any research on variations and determinants of use of medical procedures will be seriously compromised. Since 1990, diagnoses and procedures have been recorded only with specific check-boxes in the MBR.

Numerous strategies have been proposed to improve the data quality in large medical registries. Most authors regard clear instructions and continuous training as prerequisites for reliable data collection. A dialogue between the primary data producers and the data users has been enhanced by establishing a newsletter in conjunction with some national registries. In the long run, productive high quality use of data, which is also made visible to the data producers, is believed to motivate them to produce better data.

In spite of the changes and improvements that have been implemented, there seem to be few reports assessing these measures. It is strongly recommended that the effect of these efforts on data quality be reported. Those in charge of developing the registries need a firm basis for their decisions.

There is, inevitably, a limit beyond which it is more or less impossible to achieve further improvement in data quality of large nationwide registries. The recent advances in data processing technology have made it possible to introduce local information systems cost effectively into the clinical environment. One good example, a local on-line maternity data collection system, has been described by Maresh et al.[28] The system served as a data base during the stay in hospital, and as a means of communication between the clinic and community midwives. In a database that was obviously perceived as an integral part of quality conscious clinical work, a data validity level was achieved that has only been dreamt of in nationwide registries.

These experiences suggest two alternative strategies for the long term development of nationwide registries. These could be streamlined by removing the unnecessary duplication of data content already covered more reliably by local data bases. They should include only data which are obtainable reliably with relative ease, and the use of which requires a nationwide approach. In practice, this would lead to giving emphasis to administrative planning applications, and research on the use of health services and some 'hard' outcome measures.

The other strategy would rely on considerable investment in developing and standardising local databases. Extracting the chosen items from the local databases for recording in a nationwide registry would be simple in technical terms. This line of development would both strengthen the local tools for quality assurance of any health care programme and make a new level of data quality attainable in national registries.

1 McCarthy BJ, Terry J, Rochat RW, Quave S, Tyler CW. The underegistration of neonatal deaths: Georgia 1974–77. *Am J Public Health* 1980; **70**: 977–82.
2 Hammes LM, Treloar AE. Gestational interval from vital records. *Am J Public Health* 1970; **60**: 1496–1505.
3 Greb AE, Pauli RM, Kirby RS. Accuracy of fetal death reports: comparison with data from an independent stillbirth assessment program. *Am J Public Health* 1987; **77**: 1202–6.
4 Teperi J, Hemminki E, Meriläinen J, Savonen E. The reduction of distortions in perinatal statistics with a record linkage scheme. *European Journal of Public Health* 1991; **1**: 68–74.
5 Hemminki E, Teperi J, Tuominen K. Need for and influence of feedback from the Finnish birth register to data providers. *Quality Assurance in Health Care* 1992; **4**: 133–139.
6 Stephenson PA. *International differences in the use of obstetrical interventions.* Copenhagen: World Health Organisation, 1992; EUR/ICP/MCH 112.
7 David RJ. The quality and completeness of birthweight and gestational age data in computerized birth files. *Am J Public Health* 1980; **70**: 964–73.
8 Boyd NF, Pater JL, Ginsburg AD, Myers RE. Observer variation in the classification of information from medical records. *Journal of Chronic Diseases* 1979; **32**: 327–32.
9 Demlo LK, Campbell PM, Spaght Brown S. Reliability of information abstracted from patients' medical records. *Med Care* 1978; **12**: 995–1005.

10 Horwitz RI, Yu EC. Assessing the reliability of epidemiologic data obtained from medical records. *Journal of Chronic Diseases* 1984; **11**: 825–31.

11 Cameron HM, McCoogan E. A prospective study of 1152 hospital autopsies: I. Inaccuracies in death certification. *J Pathol* 1981; **133**: 273–83.

12 Cole SK. Accuracy of death certificates in neonatal deaths. *Community Medicine* 1989; **11**: 1–8.

13 Hewson D, Bennett A. Childbirth research data: medical reports or women's reports. *Am J Epidemiol* 1987; **125**: 484–91.

14 Kosecoff J, Fink A, Brook R, Chassin MR. The appropriateness of using a medical procedure—is information in the medical record valid? *Med Care* 1987; **25**: 196–201.

15 Romm FJ, Putnam SM. The validity of the medical record. *Med Care* 1981; **19**: 310–5.

16 Lilienfeld AM, Parkhurst E, Patton R, Schlesinger ER. Accuracy of supplemental medical information on birth certificates. *Public Health Rep* 1951; **66**: 191–8.

17 Querec LJ. *Comparability between the birth certificate and the national natality survey.* Hyattsville, Maryland: National Center for Health Statistics, 1980; DHEW publication no (PHS)80–1357. (Vital and health statistics; series 2; no 83).

18 Roos LL, Sharp SM, Wajda A. Assessing data quality: a computerized approach. *Soc Sci Med* 1989; **28**: 175–82.

19 Frost F, Starzyk P, George S, McLaughlin JF: Birth complication reporting: the effect of birth certificate design. *Am J Public Health* 1984; **74**: 505–6.

20 Cnattingius S, Ericson A, Gunnarskog J, Källen B. A quality study of a medical birth registry. *Scand J Soc Med* 1990; **18**: 143–148.

21 Barry J. An evaluation of the notification of birth form. *Ir J Med Sci* 1989; **158**: 102–4.

22 Carucci PM. *Reliability of statistical and medical information reported on birth and death certificates.* Albany, New York: New York State Department of Health, 1979: Monograph no 15.

23 Lumley J, Baskin S, Rigoni S. *Rights and wrongs: a validation study of the perinatal morbidity form.* Melbourne, Victoria: Victorian Perinatal Data Collection Unit, 1982.

24 Floyd L, Lavoie M, Terry JS. The status of birth certificate information in Georgia. *J Med Assoc Ga* 1981; **70**: 871–3.

25 Fedrick J, Butler NR. Accuracy of registered causes of neonatal deaths in 1958. *British Journal of Preventive and Social Medicine* (London) 1972; **26**: 101–5.

26 Gregg JB, Stanage WF, Johnson W. Birth certificate data: How reliable? *S D J Med* 1984; **37**: 21–2.

27 MacCubbin PA, Carucci PM, Tarmey JJ. Comparison of fetal death certificates and hospital records. *New York Quarterly Vital Statistics Review* 1982; 17–25.

28 Maresh M, Dawson AM, Beard RW. Assessment of an on-line computerized perinatal data collection and information system. *Br J Obst Gynaecol* 1986; **93**: 1239–45.