Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Validation of spontaneous abortion diagnoses in the Danish National Registry of Patients

This article was published in the following Dove Press journal:
Clinical Epidemiology
22 October 2010
Number of times this article has been viewed

Sarah Rytter Lohse[1]
Dóra Körmendiné Farkas[1]
Nicolai Lohse[1]
Sven Olaf Skouby[2]
Finn Erland Nielsen[3]
Timothy L Lash[1]
Vera Ehrenstein[1]

[1]Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus N, Denmark; [2]Department of Obstetrics and Gynecology, [3]Department of Cardiology, Herlev University Hospital, Herlev, Denmark

Purpose: The purpose of this study is to validate the diagnosis of spontaneous abortion (SA) recorded in the Danish National Registry of Patients (DNRP).

Methods: We randomly selected patients registered in the DNRP with a diagnosis of SA between 1980 and 2008 from hospitals in the county of North Jutland and searched for their discharge records in hospital files. We estimated positive predictive value (PPV) of the DNRP diagnosis and stratified the analysis by period (1980–1994 versus 1995–2008), hospital type (regional versus local), and International Classification of Diseases revisions (ICD-8 versus ICD-10).

Results: We could identify hospital files of 117/174 (67%) sampled registration records. Of those, the diagnosis was confirmed in 114 patients, yielding a PPV of 97.4% (95% confidence interval = 92.7%–99.5%). The PPV did not markedly vary by period, hospital type, or ICD revision. Among the three patients with available data who did not fulfill the criteria for SA, one had an induced abortion and two had threatened abortion but did not miscarry.

Conclusion: Registration of SA in the DNRP accurately reflects the diagnoses recorded in medical charts. The DNRP is a suitable source of data on SAs for epidemiologic research.

Keywords: spontaneous abortion, validation, diagnosis, Danish National Registry of Patients, positive predictive value

Introduction

Spontaneous abortion (SA) is defined as spontaneous onset of labor and evacuation of the fetus before the fetus is considered viable.[1] In Denmark, the viability threshold is set at 22 weeks of gestation. SA is an undesired pregnancy outcome that merits studies on its etiology and prognosis. Denmark is privileged by having a broad network of population-based registries,[2] which enables use of routinely collected data to rapidly assemble large cohorts for epidemiologic studies, including studies of SA. Validity of results of registry-based studies on SA depends, among other things, on the ability to reliably identify women with SA from registry records.

The true prevalence of SA in the population is difficult to estimate. In one Danish study, self-reported prevalence of SA in pregnancies intended to be carried to term was 21%, against a corresponding proportion of 16% based on Danish National Registry of Patients (DNRP).[3] Cases of SA that were not hospitalized remain undetected in registry-based studies. Based on three Scandinavian studies, 17%–30% of self-reported SAs were not registered.[3–5] Using self-report as a measure of true prevalence, however, may be questionable owing to recall bias, especially in retrospective studies, in which time since the last SA relative to self-report and early gestational age at SA were both associated with higher risk of reporting errors.[6]

Correspondence: Sarah Rytter Lohse
Department of Clinical Epidemiology, Aarhus University Hospital, Olof Palmes Allé 43-45, 8200 Aarhus N, Denmark
Tel +4589424800
Email sarahlohse@dadlnet.dk

The DNRP[7] has population-based coverage in a universal health care setting, which diminishes the risk of selection and recall bias in studies based on the registry's records. Misclassification bias, however, cannot be ruled out. At the same time, in a given study, coding errors and diagnostic underascertainment in the DNRP can be assumed to be independent of the other variables of interest, particularly if the latter is drawn from other registries. In the absence of false-positive registration, imperfect sensitivity of the SA registration, no matter how low, will leave relative estimates of association unbiased[8] if SA is the outcome of interest.

Validity of diagnoses registered in the DNRP was found to vary according to type of medical service rendered and diagnosis in question. Positive predictive values (PPVs) of surgical procedures and associated diagnoses tend to be higher than PPVs of medical diagnoses,[9] but PPVs of the latter vary greatly.[10–12] To our knowledge, no study has validated DNRP codes of SA. We aimed to validate the DNRP diagnosis of SA by estimating the proportion of SA diagnoses registered in the DNRP that could be confirmed by hospital discharge records. In addition, we aimed to determine whether the validity of the SA diagnosis varied by the revision of the International Classification of Diseases (ICD) used to code the diagnoses,[9] by hospital type, or by calendar period.

Methods

Data sources

DNRP

Denmark has 5.3 million inhabitants, and the National Health Service provides universal tax-supported health care. The DNRP tracks each hospital admission in Denmark since 1977, recording dates of admission and discharge and up to 20 discharge diagnoses. The registry covers 99.4% of all discharge records from Danish hospitals.[7] Diagnoses are coded by physicians using the Danish version of the International Classification of Diseases, 8th revision (ICD-8) (1977–1993)[13] and 10th revision (1994 onward).[14] Since 1995, outpatient contacts and visits to the emergency room have been recorded in addition to the hospital stays.

Discharge records

Discharge records were obtained from hospital files in the former county of North Jutland with 578,839 inhabitants or 10% of the population of Denmark in 2008.[15] From 1995 onward, all files have been stored electronically. Restructuring of the national health care system in the recent years caused some small local hospitals to close, and their archives are no longer readily accessible. Therefore, nonelectronic records covering the period before 1995 were available only for some of the county's small hospitals. For the large regional hospital (the hospital of Aalborg), all records were available.

Data linkage

The unambiguous individual-level linkage between the DNRP and the hospital records was enabled by the Danish unique personal 10-digit identifier (the CPR number), assigned by the Central Personal Registry at birth or immigration, encoding date of birth and sex.[16]

Selection of study population and verification of SA diagnoses

In DNRP, we identified women hospitalized in the former county of North Jutland with the diagnosis of SA between 1980 and 2008 (29 years). The ICD-8 codes were 6346x (missed abortion), 6451x (missed abortion with dilation and curettage), 6438x (SA, other), and 6439x (SA with complications); the ICD-10 codes were O021 + O021A (missed abortion) and O03x (SA). Because missed abortion and SA describe two different presentations of the same condition, ie, a nonviable fetus, we have coded the two together. In our study population, we randomly selected six patients in each calendar year (174 patients) using the SAS RANUNI function and used CPR numbers to search for corresponding discharge records. A medical doctor (SRL) reviewed the discharge records and decided for each record whether an SA diagnosis could be established by looking at the description of symptoms, observations made by physicians, blood samples, and ultrasound descriptions. Gestational age at the time of SA was obtained from discharge records.

Statistical analysis

We obtained the distribution of ICD codes of SA diagnosis and computed age of the women in the sample. The main outcome was PPV, defined as the proportion of patients with SA diagnosis in the DNRP who had the SA diagnosis verified by discharge record review. We calculated PPV with exact 95% confidence intervals (CI). PPV was calculated for the whole study population and stratified by calendar period: from 1980 to 1994 and from 1995 to 2008. This stratification was done because of the difference in the proportion of discharge records that could be ascertained in the two periods owing to the record computerization in 1995. In addition, PPVs were calculated according to hospital type (regional versus local) and ICD revision (ICD-8 versus ICD-10).

We analyzed data with Stata software (version 10.0; Stata Corp, College Station, TX). The study was approved by the Danish Data Protection Agency journal no. 2010-41-4690.

Results

Description of study population

The distribution of the 174 registered SA diagnoses is shown in Table 1. The median age for the women diagnosed with an SA was 29.0, ranging from 16.1 to 45.8 years. Discharge records could be obtained from all nine hospitals in the county of North Jutland, Denmark. Of the 174 records, 88 were admissions between 1980 and 1994, and 86 were admissions between 1995 and 2008. For the 88 earlier DNRP records, only 39 (44%) discharge records were available, whereas 78 of the 86 (91%) in the second group could be accessed. Mean and median (range) age of the 117 women with available records were 29.1 and 29 (16–45) years, respectively, and those of the 57 women without available records were 29.1 and 28 (17–45) years, respectively. Gestational age was available from 107 of the 117 records. Median (range) gestational age was 11 (6–22) weeks.

Validation of SA diagnosis

Of the 117 registry records with available discharge summary data, 114 had SA diagnosis, PPV = 97.4% (95% CI: 92.7%–99.5%). Among the DNRP records from 1980 to 1994, SA diagnosis was confirmed by discharge record review in 37 patients out of 39, corresponding to a PPV of 94.9% (95% CI: 82.7%–99.4%). Among the DNRP records from 1994 to 2008, 77 of the 78 records were verified by review, rendering a PPV of 98.7% (95% CI: 93.1%–100%). Among the three patients whose DNRP registration diagnosis could not be confirmed, one had an induced abortion and the remaining two were under observation for SA, but a living fetus was found. The two patients under observation for SA did not appear later in the DNRP with an SA or an extrauterine pregnancy within 1 year from time of registered SA. The PPVs in strata of period, hospital type, and ICD revision showed slight difference from the overall PPV (Table 2).

Table 1 Distribution of diagnosis codes

Diagnosis code	Diagnosis	No. of cases sampled
All codes		**174**
ICD-8 codes		**84**
6346x	Missed abortion	3
6451x	Missed abortion with D + C	13
6438x	Spontaneous abortion, other	6
6439x	Spontaneous abortion with complications	62
ICD-10 codes		**90**
O021 + A	Missed abortion	45
O03x	Spontaneous abortion	45

Abbreviation: ICD, International Classification of Diseases.

Discussion

We found that an overwhelming majority of SA diagnoses registered in the DNRP could be confirmed by the medical records and detected little variation in the PPV by calendar period, hospital type, or type of disease classification. Although we examined only a segment of the national population, the results can be assumed to be generalizable, thanks to the uniform structure of record keeping and the universal nature of the health care system across the entire nation. That some records were unavailable could have introduced bias, but we have no reason to believe that coding practices in those hospitals that were closed down differed from those in other hospitals, and the ages of the women were similar in the two groups. For logistical reasons, we could not assemble a sample of women based on their true SA status as reflected in medical records. Therefore, our study does not allow for estimation of either sensitivity or specificity of the DNRP diagnosis, which is a limitation. We can imagine four scenarios in which DNRP would be unable to capture cases of SA. First, hospitalized cases of SA may not be registered in DNRP because they are not coded correctly due to human or mechanical errors as described in Nickelsen's article.[9] Second, women with SA seeking medical care may not be hospitalized and thereby registered in DNRP. This circumstance is likely to be rare because all pregnant women presenting with symptoms of SA are referred to a hospital for an ultrasound examination to verify the diagnosis. Third, a woman who does not seek medical care for her SA will not be hospitalized and therefore not registered in the DNRP. The three Scandinavian studies showing discrepancy between self-reported SA and register-identified SA showed this to be a common scenario.[3–5] Finally, SA may remain unnoticed by women who are unaware of being pregnant and mistake symptoms of SA for heavy delayed menstruation. At the same time, if SA is used as the outcome of interest in epidemiologic studies, imperfect sensitivity, no matter how low, is not expected to bias relative measures of association as long as the specificity is 100%. We showed that SA cases that are registered largely represent true instances of this event. The high PPV estimated in this study indicates a low false-positive rate of DNRP diagnosis, suggesting that the DNRP is a valid source for identifying cases of SA for

Table 2 Validity of spontaneous abortion diagnosis

	No. of cases sampled	No. of records available	No. of confirmed diagnosis	PPV (%)	95% CI
All	174	117	114	97.4	92.7%–99.5%
Period					
1980–1994	88	39	37	94.9	82.7%–99.4%
1995–2008	86	78	77	98.7	93.1%–100.0%
Hospital type					
Aalborg hospital	97	91	89	97.8	92.3%–99.7%
Local hospitals	77	26	25	96.2	80.4%–99.9%
ICD revision					
ICD-8	84	36	34	94.4	81.3%–99.3%
ICD-10	90	81	80	98.8	93.3%–100.0%

Abbreviations: PPV, positive predictive value; ICD, International Classification of Diseases; CI, confidence interval.

epidemiologic research, particularly for studies that estimate ratio measures of association. Estimates of the risk of SA are likely to be underestimated.

Acknowledgments

We are very grateful to Elisabeth Kristoffersen who helped retrieve and systemize discharge records from the county of North Jutland. We also thank Anders Riis for sampling the DNRP. The study was supported by a grant provided by the Hospital of Herlev Research Committee. We thank Jacqui Goodall for linguistic revision of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended 1976 Oct 14. *Acta Obstet Gynecol Scand*. 1977;56(3):247–253.
2. Frank L. Epidemiology. When an entire country is a cohort. *Science*. 2000;287(5462):2398–2399.
3. Buss L, Tolstrup J, Munk C, et al. Spontaneous abortion: a prospective cohort study of younger women from the general population in Denmark. Validation, occurrence and risk determinants. *Acta Obstet Gynecol Scand*. 2006;85(4):467–475.
4. Axelsson G. Use of questionnaires in a study of spontaneous abortion in a general population. *J Epidemiol Community Health*. 1990;44(3):202–204.
5. Lindbohm ML, Hemminki K. Nationwide data base on medically diagnosed spontaneous abortions in Finland. *Int J Epidemiol*. 1988;17(3):568–573.
6. Wilcox AJ, Horney LF. Accuracy of spontaneous abortion recall. *Am J Epidemiol*. 1984;120(5):727–733.
7. Andersen TF, Madsen M, Jorgensen J, Mellemkjoer L, Olsen JH. The Danish National Hospital Register. A valuable source of data for modern health sciences. *Dan Med Bull*. 1999;46(3):263–268.
8. Brenner H, Savitz DA, Gefeller O. The effects of joint misclassification of exposure and disease on epidemiologic measures of association. *J Clin Epidemiol*. 1993;46(10):1195–1202.
9. Nickelsen TN. Data validity and coverage in the Danish National Health Registry. A literature review. *Ugeskr Laeger*. 2001;164(1):33–37.
10. Madsen M, Davidsen M, Rasmussen S, Abildstrom SZ, Osler M. The validity of the diagnosis of acute myocardial infarction in routine statistics: a comparison of mortality and hospital discharge data with the Danish MONICA registry. *J Clin Epidemiol*. 2003;56(2):124–130.
11. Vestberg K, Thulstrup AM, Sorensen HT, Ottesen P, Sabroe S, Vilstrup H. Data quality of administratively collected hospital discharge data for liver cirrhosis epidemiology. *J Med Syst*. 1997;21(1):11–20.
12. Kjaergaard J, Clemmensen IH, Thomsen BL, Storm HH. Validity of diagnoses of and operations for nonmalignant gynecological conditions in the Danish National Hospital Registry. *J Clin Epidemiol*. 2002;55(2):137–142.
13. World Health Organization. *Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death (ICD-8)*. Geneva, Switzerland: WHO; 1967.
14. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems, 10th Revision*; 2010. Available from: http://apps.who.int/classifications/apps/icd/icd10online/. Accessed 2010 Sep 7.
15. Statistics Denmark; 2010. Available from: http://www.statistikbanken.dk/FOLK1. Accessed 2010 Sep 7.
16. Pedersen CB, Gotzsche H, Moller JO, Mortensen PB. The Danish Civil Registration System. A cohort of eight million persons. *Dan Med Bull*. 2006;53(4):441–449.