

Reliability of recurrent pregnancy loss diagnosis coding in the Swedish National Patient Register: a validation study

This article was published in the following Dove Press journal:
Clinical Epidemiology

Emma Rasmak Roepke¹
Ole Bjarne Christiansen²
Stefan R Hansson¹

¹Department of Obstetrics and Gynecology, Institute of Clinical Sciences Lund, Skåne University Hospital, Lund University, Lund, Sweden; ²Department of Obstetrics and Gynecology, Aalborg University Hospital, Clinical Institute at Aalborg University, Aalborg, Denmark

Background: The Swedish National Patient Registry (NPR) is a nationwide registry that is used extensively for epidemiological research. Using the NPR, we recently found a recurrent pregnancy loss (RPL) incidence of 650/100,000 (0.65%) pregnant women in Sweden. It is of great importance that the quality of the coding is good and reliable in order to use NPR data for research. To specifically study RPL in Sweden, a general validation of this diagnosis in the NPR is needed.

Objective: To validate the diagnosis of RPL, defined as ≥ 3 consecutive miscarriages before 22 gestational weeks, recorded in the NPR and assess how registered miscarriages were verified clinically (ultrasound or urine/serum hCG) by reviewing the medical records.

Materials and methods: In a cohort of 6,852 women diagnosed with RPL in Sweden, during 2003–2012, a total of 238 complete medical records from 38 hospitals were reviewed. A power calculation estimated that 228 medical records had to be reviewed for a positive predictive value (PPV) of 85% (95% CI) with a power of 90%. The ICD-10 diagnoses used for RPL were N96.9 and O26.2.

Results: The diagnosis of RPL was confirmed in 202 out of 238 medical records resulting in a PPV of 85% (95% CI 78–89%) out of which 59% were verified with ultrasound whereas 35% were verified only by urine/serum hCG.

Conclusion: The Swedish NPR is a valuable tool for epidemiological research. We found a high PPV of RPL in the NPR, supporting the use of these data for future research.

Keywords: abortus habitus, recurrent pregnancy loss, validation study, registries, epidemiology

Introduction

Miscarriage is one of the most common early pregnancy complications. When it happens repeated times it is not only bad luck but is considered as a disease defined as recurrent pregnancy loss (RPL).^{1–3} RPL is traditionally defined as three or more consecutive miscarriages before the fetus reaches viability, ie, before 22 gestational weeks.⁴ Recently the definition has been debated and changed by the European Society of Human Reproduction and Embryology (ESHRE) and the Practice Committee of the American Society for Reproductive (ASRM) to as few as two miscarriages.^{2,5–7}

The prevalence of RPL is stated to be 0.5–2.3%,^{1,2,8–10} and the incidence has been estimated at 650/100,000 pregnant women based on the results from a retrospective register study in Sweden.¹¹ Estimating the incidence of RPL is somewhat

Correspondence: Emma Rasmak Roepke
Skåne University Hospital, Obstetric and
gynecological department, Jan
Waldenströmsgatan 47, 205 02 Malmö,
Sweden
Tel +46 452 346 0721
Email emma_36@hotmail.com

difficult because of a lack of consensus regarding the definition of RPL but also due to uncertainty about whether to include biochemical miscarriages (positive urinary human chorionic gonadotropin (hCG), a raised serum hCG) or only clinical miscarriages verified with ultrasound or histology.^{2,3,12,13} The above incidence of RPL in Swedish women is to our knowledge the first study on the RPL incidence^{11,13} as others report an occurrence or a prevalence of PRL.^{1,9} The incidence based on registers will also depend on if women with miscarriage get in contact with a physician in the health care system.¹³ This often happens if bleeding occurs in early pregnancy or a routine scan is performed.

Women and/or couples with RPL are often in distress. However, a possible explanation for why they miscarry is only found in 40–50% of the cases. Parental chromosomal translocations, congenital and acquired uterine abnormalities, endocrine imbalances, autoimmune factors including the antiphospholipid syndrome, as well as infections and thrombophilia, are all possible causes.^{10,12,14,15}

Many epidemiological studies on RPL are observational and based on data from registers that are available for research purposes. Such studies are also important to improve the quality of health care, and to manage health care services.^{16,17}

One such register is the Swedish National Patient Register (NPR), which was used in the above incidence study on RPL.^{11,17} Clearly, it is important that the data in the registers are valid and representative. A recent review stated that the accuracy for many diseases in the NPR register ranged from 85% to 95%.¹⁶

To our knowledge, there is no evidence regarding the precision of RPL in the NPR. Such evidence is essential for the use and interpretation of RPL based on NPR data. The aim of this study was to validate the ICD-10 diagnoses codes registered in the NPR for RPL (N96.9, O26.2). A secondary aim was to analyze whether only miscarriages verified by ultrasound or also pregnancy losses verified only by urine or serum hCG are registered in the NPR and used for the RPL diagnosis.

Material and methods

Study design

A cross-sectional observational study designed was used to validate the RPL diagnosis registered in the NPR compared with the information obtained in the medical records. The medical records were considered as “reference standard”. In order to assess whether a diagnosis of

RPL was defined as three or more consecutive miscarriages, the medical records were reviewed.

The main aim was to evaluate the positive predictive value (PPV), defined as the proportion of women registered as RPL in the NPR with a diagnosis verified by information from medical records. Sensitivity was assessed and defined as the proportion of women suffering RPL and registered with the diagnosis in the NPR.

Population

In our previous study, a cohort of 6,852 women was diagnosed with RPL in the NPR over a 10-year period in Sweden.¹¹ We assumed that only one-third of the medical records requested from hospitals would be available and/or would provide the necessary information. A sample size of 711 women, registered with the ICD-10 diagnoses codes N96.9 and/or O26.2 were randomly selected by the National Board of Health and Welfare (NBHW), who are responsible for the NPR (Figure 1).

Data sources

The NPR includes information about hospitalization, recorded at discharge. Each record states admission and discharge dates, hospital department, and the primary and up to seven contributory diagnoses coded according to the Swedish versions of the current edition (at the time of diagnosis) of the International Classification of Diseases (ICD).¹⁷

NBHW is the Swedish agency that administrates and is responsible for the NPR. Reporting information to the NPR is mandatory and is delivered to the NBHW once a month from each of the 21 county councils in Sweden. Since 1987, the NPR has included all inpatient events in Sweden. Since 2001, the register has also covered outpatient visits, including both private and public caregivers. The coverage of in-patients is currently almost 100%. For hospital-based outpatients, data from some private caregivers are missing, so the total cover for outpatients is considerably lower (approximately 80%).^{16,18}

The samples selected by NBHW were received encrypted with the 10-digit personal identification number assigned to all Swedish residents, as well as a hospital and clinic code. The selected hospitals provided a variation in demographic characteristics from larger university hospitals to smaller regional hospitals with geographic representation from all regions. All relevant available clinics were contacted to obtain the medical records. Consent was secured from the

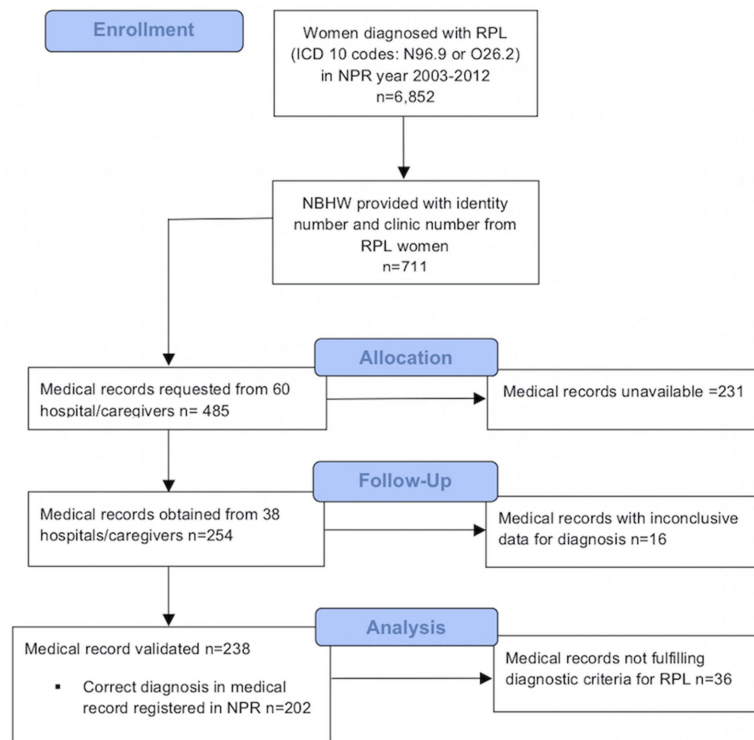


Figure 1 Flowchart of women diagnosed with recurrent pregnancy loss in the national patient register and validated with medical records.

head of each clinic. The medical secretary or archive staff at the hospitals provided the medical records by registered mail. The first author reviewed the medical records and the RPL diagnoses (N96.9 and O26.2) were compared and validated against data in the NPR.

In our previous incidence study, an estimation was made of how many women, registered with three or more consecutive miscarriages, ie, fulfilling the diagnostic criteria of RPL, lacked the registered diagnosis codes of RPL (N96.9, O26.2) in the NPR.¹¹ This group comprised women, without an RPL diagnosis in the register, with three or more of the following ICD codes for miscarriages (main or bi-diagnosis) without interspersed live births: O02.1, O03 (including sub-diagnosis O03.0–O03.9) during 2003 and 2012.¹¹ Since the same diagnosis can be used several times on different occasions when contacting the healthcare system about the same pregnancy, a 90-day interval was required between two registered diagnoses of miscarriage to count as a new miscarriage. This group of women, with three or more consecutive miscarriages fulfilling the diagnostic criteria for RPL, but without registration in the NPR as RPL, are considered “false negative” and are used for calculating the sensitivity of RPL in the NPR.

Variables used

The method of diagnosing miscarriage: ultrasound examination or diagnosis based on a positive hCG-test, as well as method of conception: spontaneous or in vitro-fertilization pregnancy, were evaluated. Furthermore, the distribution regarding primary or secondary RPL and known risk factors for RPL were also evaluated.

Statistical analysis

The accuracy of positive register diagnoses of RPL was determined by PPV, ie, the probability of positive registered cases being confirmed by the medical records (reference standard). The PPV was calculated as: [true positive register cases] divided by [true positive in reference standard] plus [false-positive register cases] with 2×2 cross tables and a χ -square test with 95% confidence intervals (95% CI).

The ability of the registers to detect women with three or more consecutive miscarriages (“true positive”) was determined by sensitivity, ie, the probability of positive cases ever appearing in the register with a diagnosis of RPL. Sensitivity was calculated as: [true positive register cases] divided by [true positive register cases] plus [false

Table 1 Validation of RPL diagnosis in medical records, randomly selected from women diagnosed with RPL in the NPR

Diagnoses in medical records	Number of cases, n	Positive predictive value % (95% CI)
Total medical records reviewed	238	
3 consecutive miscarriages	202	85 (80–89%)
Incorrect diagnoses	36	–
– 3 non-consecutive miscarriages	7	
– Only 2 miscarriages	16	
– 1 of the 3 miscarriages was a IUFD (>gw.22)	4	
– 1 of the 3 miscarriages was a termination of pregnancy because of chromo-some abnormality	1	
– 1 or 0 miscarriages	7	
– Other reason	1	

Abbreviations: RPL, recurrent pregnancy loss; ICD, International Classification of Diseases; NBHW, National Board of Health and Welfare (responsible for the NPR); NPR, national patient register; IUFD, intrauterine fetal demise; gw, gestational week.

negative register cases] with 2×2 cross tables and a χ^2 -square test with a 95% CI.

A power analysis, assuming that PPV would be 85% and the accepted width of a 95% CI was 10 percentage units (eg, 80–90%), calculated that a sample size of 228 medical records were needed to achieve a power of 90%.

All data were analyzed using SPSS version 24 (Armonk, NY: IBM Corp.)

Ethics

The Regional Ethical Board at Lund University (Lund, Sweden) approved the national study (Dnr 2016/478).

Results

In the randomly selected cohort of 711 women, registered with RPL in the NPR, 485 medical records were requested from 60 different hospitals and private caregivers. Out of the 60 hospitals, 73% of the university hospitals, 60% of the regional hospitals and 50% of the private caregivers responded. In total, 254 medical records were available for review from 38 hospitals. In 16 cases, the medical records were not conclusive regarding the correctness of the diagnoses as information was missing (Figure 1). Correct diagnosis was made in 202 out of the 238 medical records, resulting in PPV $(202/238)=85\%$ (95% CI 78–89%) (Table 1). Out of the remaining 36 women with incorrect diagnosis, 27 cases (16+4+1 (Table 1)) could have been included as correct if the definition of two miscarriages had been used to qualify as RPL.

In our previous study about RPL incidence,¹¹ a cohort consisting of 990 women had had three consecutive miscarriages during 2003–2012 but with missing RPL

diagnosis in the NPR. These women were defined as “false negative” as they fulfilled the criteria for the diagnoses of RPL but were not registered as such in the NPR. In the same study,¹¹ a cohort of 6,852 women was diagnosed with RPL in the NPR during the 10-year period. With a PPV of 85%, estimated in this current study, the number of “true positives” in the cohort registered with RPL was calculated to be 5,824 women $(0.85 \times 6,852 = 5,824)$. For calculating sensitivity, “true positive” was divided by “true positive” plus “false negative” which gave $5,824 / (5,824 + 990) = 0.85$, ie, a sensitivity of 85%.

Pregnancy variables were evaluated among the women who were correctly diagnosed ($n=202$) (Table 2). One hundred and twenty (59%) women had their miscarriage verified with ultrasound, whereas for 70 (35%) it was verified only by urine/serum hCG. There were 100 women (50%) without known risk factors for RPL. One hundred and eleven (55%) and 88 (44%) women had primary and secondary RPL, respectively.

Discussion

This validation study for RPL diagnosis registered in the Swedish NPR showed 85% PPV and a sensitivity of 85%. The findings are important for future research using data from the NPR in general and on RPL specifically. The data confirm our previous incidence study.¹¹

To our knowledge, no other studies have evaluated the data quality of registered RPL in national health registers. To avoid selection bias and confounding factors, our validation study was conducted with data from a large number of hospitals from different regions in Sweden. The responding rate for university hospitals (73%) and regional hospitals

Table 2 Women with correct diagnosis of recurrent pregnancy loss and pregnancy variables

Variables	Number of women (%) (Total n: 202)
Method for verifying pregnancy All three pregnancies verified by ultrasound One or more pregnancies only verified by hCG measurement Unknown	120 (59) 70 (35) 12 (6)
Conception method 3 spontaneous 2 spontaneous + I IVF 2 or 3 IVF	186 (92) 5 (2.5) 11 (5.5)
Possible cause of RPL Uterine malformation Uterine septae Uterine fibroma Uterine polyp Uterine synechiae Hypothyroidism Parental chromosomal translocation Antiphospholipid syndrome (positive anticardiolipin antibodies or lupus anticoagulant) Cervical insufficiency Thrombophilia None found Not examined	5 (2.5) 11 (5.5) 3 (1.5) 3 (1.5) 5 (2.5) 7 (3.5) 8 (4) 5 (2.5) 1 (0.5) 4 (2) 100 (50) 49 (24)
RPL Primary Secondary Tertiary Inconclusive	111 (55) 88 (44) 1 (0.5) 2 (1)

Abbreviations: hCG, human chorionadotropin; IVF, in vitro fertilization; RPL, recurrent pregnancy loss; Primary RPL, no children before recurrent miscarriages; Secondary RPL, childbirth before recurrent miscarriages; Tertiary RPL, recurrent pregnancy loss with childbirth after and subsequently again recurrent pregnancy loss.

(60%) was pretty similar. There was a good geographic and demographic distribution of the participating hospitals so the material of the responded hospital can be considered representative and one of the strengths of the study.

Some medical records (n=16) were not conclusive regarding whether the RPL criteria were met or not. Furthermore, there were a substantial number of medical records (n=231) that could not be obtained. As high as 37% of the hospitals did not respond to the request to participate, a factor that might skew the results if there is a difference in the accuracy of diagnosing RPL in those who responded and those who did not. There are no obvious defined characteristics in the hospitals that did not respond compared to those that did respond that could suspect an influence in the results.

Despite these minor weaknesses, the power estimation of 228 medical records was reached in the study. Only one

physician reviewed the medical records and consequently, no cross-reading of the medical records was possible. Such a cross-reading could have strengthened the study further and could also have enabled an inter-observational sub-analysis of the validated diagnoses.

The proportion of valid diagnoses, ie, PPV, in the NPR is probably higher in patients with severe as opposed to mild diseases.¹⁶ RPL is considered as a benign diagnosis, and taking this into consideration, the estimated PPV of RPL in the current study is fully acceptable and in line with the overall accuracy for several other diseases registered in the NPR.¹⁶

Sensitivity concerns how complete a register is and how good the register is at capturing cases with a disease.¹⁹ To assess the sensitivity a random population should be studied for the disease and then there should be a crosscheck to see whether individuals with the disease are recorded in the

register. Many women suffer miscarriages without contacting a hospital or another caregiver and therefore do not exist in medical records.²⁰ It was not possible to study a random population for “false negative” cases; instead, we used data from our previous study describing the incidence of RPL.¹¹ This might be considered less accurate than the first mentioned sensitivity assessment, as it was dependent on coding of miscarriages in the NPR.

Whether a woman with RPL is reported in the NPR depends on several factors, including the care-seeking behavior of an individual, access to health care, and the physician’s awareness of the disease. Hospital fees can be an issue in some countries but since they are relatively low in Sweden this factor is of less importance. The NPR coding is also used as the basis for management and financing, and some ICD codes give higher financial compensation than others. The effects of financial incitements on ICD coding have probably been underestimated.¹⁶ To our knowledge, there have not been any significant changes in financial compensation for miscarriage or RPL codes over the last few decades.

Coding in registers is also associated with errors such as not having a better code to use and typing mistakes when documenting the code.²¹ Ludvigsson et al presented three types of diagnostic coding error in medical records; I) diagnostic errors, ie, the patient received an incorrect diagnosis, II) translation errors, ie, the ICD code in the NPR is different from the code actually listed in the patient chart, and III) coding errors, ie, a faulty ICD code accompanies an otherwise correct diagnosis.¹⁶

If two miscarriages had been used to define RPL as suggested by the ESHRE and the ASRM Practice Committee,^{6,7} the PPV would have been higher. However, the definition of RPL used in Sweden during this period was three or more miscarriages. Our search in the NPR included diagnoses from both in- and outpatient records, although most women with a miscarriage or RPL are treated as outpatients. The NPR started registering outpatients in 2001 and according to the NBHW, the register is reliable for the period 2003–2005. It would, therefore, have been interesting to stratify the diagnoses by year and analyze whether any differences in the validation parameters for RPL changed over the years. However, this information was not possible to obtain and a stratification could therefore not be performed. The negative predictive value and specificity could not be estimated in this study, as the numbers of “true negatives” were unknown.

Only 59% of the pregnancies were verified with ultrasound, which is interesting considering that some scientific societies require a clinical pregnancy (verified by ultrasound or histopathology) as part of the definition.⁷ Thirty-five percent of the women with RPL had at least one of the pregnancies confirmed by hCG only and many of these would not have been given the RPL diagnosis in the opinion of the ASRM Practice Committee.⁷ This mode of confirmation is, however, acceptable according to ESHRE’s RPL 2018 guideline for pregnancies past six gestational weeks with heavy bleeding.⁶ According to this definition, the verified pregnancies used in this study are therefore correctly defined.

In half of the women who were verified with a RPL diagnosis in the medical record, no possible risk factor for RPL was found. This is in line with other studies reporting on the RPL etiology.^{15,22} Of the women, 55% and 44% had experienced primary and secondary RPL, respectively, which is also in line with previously reported data.¹¹

Conclusion

The data indicate that the obtained PPV and sensitivity for the RPL diagnosis in the Swedish NPR have high quality and are therefore useful to use for future RPL research.

Acknowledgments

We thank NBHW for helping with data from the NPR. Thanks also go to all hospitals, doctors and medical secretaries that contributed medical records. This work was supported with funding from Skåne County Council’s Research and Development Foundation.

Author contributions

All authors contributed toward data analysis, drafting and critically revising the paper, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Sugiura-Ogasawara M, Suzuki S, Ozaki Y, Katano K, Suzumori N, Kitaori T. Frequency of recurrent spontaneous abortion and its influence on further marital relationship and illness: the Okazaki Cohort Study in Japan. *J Obs Gynaecol Res*. 2013;39(1):126–131. doi:10.1111/j.1447-0756.2012.01973.x
2. Larsen EC, Christiansen OB, Kolte AM, Macklon N. New insights into mechanisms behind miscarriage. *BMC Med*. 2013;11:154. doi:10.1186/1741-7015-11-154

3. Farquharson RG, Jauniaux E, Exalto N; ESHRE Special Interest Group for Early Pregnancy (SIGEP). Updated and revised nomenclature for description of early pregnancy events. *Hum Reprod*. 2005;20(11):3008–3011. doi:10.1093/humrep/dei167
4. Regan L, Backos M, Rai R. The Investigation and Treatment of Couples with Recurrent First-trimester and Second-trimester Miscarriage. Green-top guideline No. 17; April 2011. Royal College of Obstetricians and Gynaecologists. Available from: https://www.rcog.org.uk/globalassets/documents/guidelines/gtg_17.pdf.
5. Bhattacharya S, Townend J, Bhattacharya S. Recurrent miscarriage: are three miscarriages one too many? Analysis of a Scottish population-based database of 151,021 pregnancies. *Eur J Obs Gynecol Reprod Biol*. 2010;150(1):24–27. doi:10.1016/j.ejogrb.2010.02.015
6. The ESHRE Guideline Group on RPL; Bender Atik R, Christiansen OB, et al.. ESHRE guideline: recurrent pregnancy loss. *Hum Reprod Open*. 2018;2:hoy004.
7. Practice Committee of American Society for Reproductive Medicine. Definitions of infertility and recurrent pregnancy loss: a committee opinion. *Fertil Steril*. 2013;99(1):63. doi:10.1016/j.fertnstert.2012.09.009
8. Carp H. Recurrent pregnancy loss: towards more accurate diagnosis and treatment. *IMAJ*. 2001;3:528–532.
9. Alberman E. The epidemiology of repeated abortion. In: Beard RW, Sharp F, editors. *Early Pregnancy Loss Mech Treat*. London: Springer; 1988:9–17.
10. Jauniaux E, Farquharson RG, Christiansen OB, Exalto N. Evidence-based guidelines for the investigation and medical treatment of recurrent miscarriage. *Hum Reprod*. 2006;21(9):2216–2222. doi:10.1093/humrep/del150
11. Rasmak Roepke E, Matthiesen L, Rylance R, Christiansen OB. Is the incidence of recurrent pregnancy loss increasing? A retrospective register-based study in Sweden. *Acta Obs Gynecol Scand*. 2017;96(11):1365–1372. doi:10.1111/aogs.13210
12. Kutteh WH. Novel strategies for the management of recurrent pregnancy loss. *Semin Reprod Med*. 2015;33(3):161–168. doi:10.1055/s-0035-1552586
13. Saravelos SH, Li T-C. Unexplained recurrent miscarriage: how can we explain it? *Hum Reprod*. 2012;27(7):1882–1886. doi:10.1093/humrep/des102
14. Christiansen OB. Epidemiology of recurrent pregnancy loss. In: Carp HJA, editor. *Recurrent Pregnancy Loss: Causes, Controversies and Treatment*. London: Informa Healthcare; 2007:1–13.
15. Allison JL, Schust DJ. Recurrent first trimester pregnancy loss: revised definitions and novel causes. *Curr Opin Endocrinol Diabetes Obes*. 2009;16(6):446–450. doi:10.1097/MED.0b013e3283327fc5
16. Ludvigsson JF, Andersson E, Ekblom A, et al. External review and validation of the Swedish national inpatient register. *BMC Public Health*. 2011;11:450. doi:10.1186/1471-2458-11-450
17. The National Board of Health and Welfare. Swedish official statistics [homepage on the Internet]. 2015. Available from: <https://www.socialstyrelsen.se>. Accessed April 2016.
18. Forsberg L, Jacobsson A, Nyqvist K, Heurgren M, Rydh H. Quality and content of the Patient Register. 2008;125(1). Swedish. Available from: http://www.socialstyrelsen.se/Statistik/statistik_amne/slutenvard/patientregistret.htm. Accessed October 2018.
19. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clin Epidemiol*. 2015;7:449–490. doi:10.2147/CLEP.S91125
20. Christiansen OB. Immunology and recurrent pregnancy loss – epidemiological aspects in humans. In: Chaouat G, Sandra O, Lédée N, editors. *Immunology and Pregnancy*. Paris: Bentham eBooks; 2013:676–690.
21. Sundholm A, Burkill S, Sveinsson O, Piehl F, Bahmanyar S, Nilsson Remahl AIM. Population-based incidence and clinical characteristics of idiopathic intracranial hypertension. *Acta Neurol Scand*. 2017;136(5):427–433. doi:10.1111/ane.12742
22. Yang CJ, Stone P, Stewart AW. The epidemiology of recurrent miscarriage: a descriptive study of 1214 prepregnant women with recurrent miscarriage. *Aust N Z J Obs Gynaecol*. 2006;46(4):316–322. doi:10.1111/j.1479-828X.2006.00599.x

Clinical Epidemiology

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification,

systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: <https://www.dovepress.com/clinical-epidemiology-journal>

Dovepress