The final version of this paper was published in the *Aust NZ J Pub Health* 2014: 38:258-64.

Linking birth records to hospital admission records enhances the identification of women who smoke during pregnancy

Authors, qualifications and contact details

Duong Thuy Tran BMed, MIPH, PhD
Research Fellow
Centre for Health Research - School of Medicine
University of Western Sydney
Locked Bag 1797
Penrith NSW 2751 Australia
Email: D.Tran@uws.edu.au

Professor Christine Roberts MBBS, MPH, FAFPHM, PhD
Research Director
Clinical and Population Perinatal Health Research
Kolling Institute of Medical Research
Building B52-Royal North Shore Hospital
St Leonards NSW 2065 Australia
Email: Christine.Roberts@sydney.edu.au

Alys Havard BPsych (Hons), MPH, PhD
Research Fellow
Centre for Health Research - School of Medicine
University of Western Sydney
Locked Bag 1797
Penrith NSW 2751 Australia
Email: A.Havard@uws.edu.au

Professor Louisa R Jorm BVSc (Hons 1), MSc (Epidemiology), PhD
Director, Centre for Health Research
Professor of Population Health, School of Medicine
University of Western Sydney
Locked Bag 1797
Penrith NSW 2751 Australia
Email: L.Jorm@uws.edu.au

Principal Scientist
The Sax Institute
PO Box K617
Haymarket NSW 1240 Australia

Abstract

Objective: Birth records and hospital admission records are valuable for research on maternal smoking, but individually are known to under-estimate smokers. This study investigated the extent to which combining data from these records enhances the identification of pregnant smokers, and whether this has impact on research findings such as estimates of maternal smoking prevalence and risk of adverse pregnancy outcomes associated with smoking.

Methods: 846,039 birth records in New South Wales, Australia (2001-2010) were linked to hospital admission records (delivery and antenatal). Algorithm 1 combined data from birth and delivery admission records, whereas algorithm 2 combined data from birth record, delivery and antenatal admission records. Associations between smoking and placental abruption, preterm birth, stillbirth, and low birthweight were assessed using multivariable logistic regression.

Results: Algorithm 1 identified 127,612 smokers (smoking prevalence 15.1%), which was a 9.6% and 54.6% increase over the unenhanced identification from birth records alone (prevalence 13.8%), and delivery admission records alone (prevalence 9.8%), respectively. Algorithm 2 identified a further 2,408 smokers from antenatal admission records. The enhancement varied by maternal socio-demographic characteristics (age, marital status, country of birth, socio-economic status), obstetric factors (multi-fetal pregnancy, diabetes, hypertension), and maternity hospital. Enhanced and unenhanced identification methods yielded similar odds ratios for placental abruption, preterm birth, stillbirth, and low birthweight.

Conclusions: Use of linked data improved the identification of pregnant smokers. Studies relying on a single data source should adjust for the under-ascertainment of smokers among certain obstetric populations.

Keywords: health record linkage, smoking, birth, perinatal, hospital admission

BACKGROUND

Maternal smoking during pregnancy is one of the most preventable causes of miscarriage, fetal growth restriction, placental abruption, preterm delivery, stillbirth, and perinatal death.[1-4] However, less than half of smoking women (4%-47%) quit spontaneously when they become pregnant.[5] Between 69% and 75% of women who smoked during their first pregnancy continue to do so in subsequent pregnancies, even among those who experience stillbirth or infant death.[6-9] Despite a decline over the past two decades,[1,10,11] the prevalence of maternal smoking remains high in high-income western countries (11%-17%).[8,10-13]

In Australia, tobacco control policies, including increased taxation, smoke-free environment legislation, expanded graphic health warnings, restricted tobacco advertising, and plain packaging and display have been successful in reducing the smoking rate in the general population.[14] There have been concerted efforts among health professionals across disciplines to identify pregnant smokers and provide these women with smoking cessation advice and interventions.[2,15,16] Evaluating the effectiveness of these initiatives among obstetric populations requires population-wide measures of pregnant smokers.

Routinely collected health data, such as birth records, records of hospital admissions, medical and pharmaceutical claims, and birth and death registrations are valuable resources for research on smoking during pregnancy because of their whole-of-population coverage.[17-19] In Australia, routinely collected birth data are currently used for routine reporting and monitoring of smoking during pregnancy.[20,21] Under-reporting of smoking status, however, may bias the findings of research that relies on a single source of routine data.[22-26] Using biochemical testing to validate self-reporting of smoking status, studies conducted in Australia and elsewhere have reported that between 6% and 39% of smoking women do not disclose their smoking status.[27-31] Greater levels of non-disclosure are found in women who are older, married, and have better socio-economic status.[23,26,27] As a result, the true prevalence of maternal smoking is likely to be under-estimated and opportunities for smoking cessation interventions could be missed. Using medical records as the gold standard, Australian validation studies of birth and hospital admission data (for a delivery) showed that when smoking is reported in these datasets it is accurate (positive predictive values 96.1% and 93.0%, respectively) but it is under-ascertained (sensitivities 89.6% and 66.3%, respectively).[24,25] Such evidence indicates that these data sources, when used alone, are likely to under-estimate smoking women.

Given that biochemical testing has not yet been used as a global screening method for maternal smoking in Australia, combining information from more than one source of data is an alternative strategy to improve the identification of smokers. Linking birth data to hospital admission data has provided greater sensitivities of maternal morbidity diagnoses, birth outcomes, and intrapartum procedures without any increase in false positives.[32-36] Studies in the United States have demonstrated that combining information from birth certificates and confidential questionnaires yielded higher estimates of maternal smoking prevalence.[22,26] Little is known in Australia about the extent to which linking birth data with hospital admission data improves the identification of pregnant smokers, and its potential impact on estimates of maternal smoking prevalence, and risks of smoking-related adverse pregnancy outcomes including placental abruption, preterm delivery, stillbirth, and low birthweight. Our study addressed this information gap.

METHODS

Data sources and linkage

This study was part of a larger research project (The Smoking MUMS Study),[19] which uses birth records linked to a number of other routine data collections to investigate the utilisation and safety of smoking cessation pharmacotherapies among obstetric populations. For the current analyses, population-based birth records (2001-2010) for women residing in New South Wales (NSW), Australia were linked to records of admissions to NSW hospitals (2000-2010). Birth records were extracted from the NSW Perinatal Data Collection (referred to as '*birth record*'), which is a legislated and population-based surveillance system covering all births (live births and stillbirths of at least 20 weeks gestation or at least 400 grams birth weight) in NSW public and private hospitals, and at home. The birth data collect information on the health of mothers and babies, including maternal socio-demographic characteristics, smoking status, pre-existing and gestational medical diagnoses, and outcomes of the pregnancy.[20]

Hospital admission records were extracted from the NSW Admitted Patient Data Collection, which is a mandatory data collection of all hospital discharges from all public, private and repatriation hospitals, and private day procedure centres in NSW. The hospital admission data contains maternal socio-demographic characteristics, diagnoses, and procedures undertaken. For each hospital admission record, between one and 55 diagnoses are coded according to the International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM).[37]

Records from the two sources of data were linked by the NSW Centre for Health Record Linkage, using a probabilistic matching method[38,39] and the privacy preserving approach.[18] The validity of the probabilistic record linkage is extremely high, with less than 0.3% false positive links, and less than 0.5% missed links.[38,39] Only de-identified data were provided to researchers. Because this study identified women who smoked during pregnancy, hospital admission records were limited to the admission for the delivery (referred to as '*delivery admission*') and other admissions during the pregnancy period (referred to as '*antenatal admission*'). Birth records which were not linked to a corresponding delivery admission record were excluded (n=9,758, 1.1%). Multi-fetal pregnancies resulted in more than one birth record containing the same maternal information, for those women the first record was selected.

Study variables

The birth data include a check box ('Yes', ' No') for the item '*Did the mother smoke cigarettes at all during pregnancy?*'.[20] In hospital admission data, the assignment of the Z72.0 diagnostic code in any diagnosis field indicates the use of tobacco in the last month.[37] This study identified a pregnant smoker using the following methods:

1. Unenhanced identification from birth records alone: Smokers were identified solely from the birth records.
2. Unenhanced identification from delivery admission records alone: Smokers were identified solely from the delivery admission records.
3. Enhanced algorithm 1: A woman was identified as a smoker if this was indicated in the birth record or in the delivery admission record.
4. Enhanced algorithm 2: A woman was identified as a smoker if this was indicated in the birth record, the delivery admission record, or in the antenatal admission record.

Maternal socio-demographic characteristics included age, country of birth, marital status, residential remoteness, residential socio-economic status, type of maternity hospital, and year of delivery. Country of birth was grouped as English-speaking (Australia, New Zealand, United Kingdom and North America), Asia (South East, North East and Southern Asia), and other non-English speaking countries.[40] Socio-economic status of residential postcodes was measured by the Socio-Economic Index of Area (SEIFA, based on the Index of Relative Socio-economic Disadvantage scores)[41] and grouped into score quintiles.

Obstetric characteristics were parity, plurality, diabetes (gestational and pre-existing), and hypertension (chronic and pregnancy-induced). Pregnancy outcomes included preterm birth (<37 weeks, medically indicated or spontaneous), stillbirth, placental abruption, and small-for-gestational-age (SGA, <10th birth weight percentile). Maternal smoking is an established risk factor for these pregnancy outcomes.[1,36,42-48]

Analyses

Enhancement in identification of pregnant smokers

The number of smokers identified by the unenhanced and enhanced methods was tabulated according to maternal socio-demographic characteristics, obstetric characteristics, and pregnancy outcomes. The enhancement was presented as the percentage increase in the identification. The percentage increase when comparing the number of smokers identified by algorithm 1 with the number of smokers identified from birth records alone, for example, was calculated as below:

$$\frac{\text{N smokers identified by algorithm 1} - \text{N smokers identified from birth records alone}}{\text{N smokers identified from birth records alone}} \times 100\%$$

Impact of enhanced identification

The impact of the enhanced identification of smokers on both absolute and ratio measures was examined. To determine the impact on absolute measures, the prevalence of maternal smoking for each year between 2001 and 2010 was estimated and compared with the expected rates of smoking in pregnant women. The expected rate was interpolated using (i) the number of pregnancies in each calendar year tabulated in this study, (ii) annual rate of female smoking (≥16 years) reported by NSW Population Health Surveys (PHS, population-based phone-interviews, 2002 to 2010),[49] and (iii) an overall 4% quitting rate among pregnant women in NSW.[50]

To examine the impact of the enhanced identification on ratio measures, the risk of preterm birth, stillbirth, placental abruption, and SGA associated with maternal smoking was assessed using multivariable logistic regression. The odds ratios (OR) were adjusted for year of delivery, maternal age, country of birth, marital status, SEIFA, parity, diabetes, and hypertensive disorders. Only singleton pregnancies (n=832,859) were included in the logistic regression analyses. It was considered whether adjusted ORs for preterm birth, stillbirth, placental abruption, and SGA were consistent with findings from other Australian and international population-based studies that used routinely collected data sources or longitudinal cohort data.[1,42-48]

All analyses were carried out in SAS version 9.3. Ethical approval for this study was granted by the NSW Population and Health Services Research Ethics Committee.

RESULTS

Enhanced identification of smokers

The final datasets included 846,039 birth records and 1,189,612 hospital admission records (846,039 delivery admissions and 343,573 antenatal admissions) for 554,097 women. The mean maternal age was 30.4 years (standard deviation 5.6.years, median 30.7 years). Table 1 shows the number of smokers identified by the enhanced algorithm 1 and percentage increase over the identification by the unenhanced methods. In the entire linked dataset, 127,612 smokers were detected using the enhanced algorithm 1, whereas 116,387 and 82,565 smokers respectively were identified from birth records alone, and delivery admission records alone.

Compared to the unenhanced identification from the birth records, algorithm 1 increased the detection of smokers by 9.6% overall. The largest increases were among women who had obstetric care in a private hospital (28.6%), who were born in a non-English speaking country (Asia 21.1%, other country 17.0%), and who lived in areas of highest socio-economic status (first SEIFA quintile 18.5%). The enhancement was around 14% among women who had a multi-fetal pregnancy, diabetes, hypertension, and a large-for-gestational-age infant. The increase was about 12% among women who were married or aged ≥30 years. A small increase (3.6%) was found among rural hospitals (Table 1).

Compared to the unenhanced identification from the delivery admission records, algorithm 1 increased the identification of smokers by 54.6% overall. The greatest increase (115.7%) was among women who were born in Asian countries (2,293 smokers identified by algorithm 1 versus 1,063 smokers identified from the delivery admission records alone). Large increases were also among those who had obstetric care in a private (96.3%) or rural (83.7%) hospital, who born in another non-English speaking country (78.2%), and who lived in the first SEIFA quintile areas (77.2%) (Table 1).

Compared with algorithm 1, the enhanced algorithm 2 identified a further 2,408 (1.9%) smokers from antenatal admission records. The increase varied only slightly according to maternal characteristics and pregnancy outcomes, with the largest increases in women who had triplets or quadruplet births (6.9%), twin births (4.5%), or who had obstetric care in a private hospital (4.5%) (Table 1).

Insert Table 1 about here

Prevalence of maternal smoking

The overall prevalence of maternal smoking estimated from delivery admission records alone was 9.8% (95% confidence interval [95%CI] 9.7-9.8), from birth records alone was 13.7% (95%CI 13.7-13.8), from algorithm 1 was 15.1% (95%CI 15.0-15.2), and from algorithm 2 was 15.4% (95%CI 15.3-15.4). Between 2001 and 2010, there was a declining trend in the prevalence of smoking in pregnant women and in the general female population (Figure 1). The prevalence of maternal smoking estimated using the algorithm 1 was comparable to the expected rates of maternal smoking for the following years: 2002 (18.2% vs 18.9%), 2006 (15.0% vs 15.6%), 2007 (13.9% vs 14.8%), 2009 (13.1% vs 13.6%), and 2010 (12.3% vs 13.0%).

Insert Figure 1 about here

Effects of smoking on pregnancy outcomes

Results of multivariable logistic regression analyses are presented in Table 2. Enhanced and unenhanced identification of smokers yielded similar adjusted ORs for placental abruption (1.94 to 2.07), preterm birth (1.59 to 1.73), stillbirth (1.21 to 1.40), and SGA (2.28-2.46). These estimates were consistent with those reported in other population-based studies conducted in Australia and elsewhere (Table 2).

Insert Table 2 about here

DISCUSSION

Routinely collected birth and hospital admission data are increasingly in use for surveillance and health research, however, these individual datasets under-estimate maternal smoking.[24,25] This study showed that combining smoking information from the birth and hospital admission data enhanced the identification of pregnant smokers and improved the estimates of prevalence of smoking during pregnancy among Australian women. Estimates of the effects of maternal smoking on placental abruption, preterm birth, stillbirth, and low birth weight were similar to those in the published literature, irrespective of whether unenhanced or enhanced methods were used.

Previous studies have reported that the accuracy and completeness of the recording of pre-existing and pregnancy-induced medical diagnoses in Australian hospital morbidity data was generally better than in birth data.[33,35] In contrast, this study found that birth records provided a higher enumeration of pregnant smokers than hospital admission records. This is likely to reflect the fact that birth records are collected primarily for surveillance purposes, and include a specific, mandatory data item for smoking. On the contrary, hospital morbidity data are collected mainly for administrative purposes, do not contain a smoking-specific item, and the recording of smoking as a diagnosis is required only when smoking affects patient management.[37] In addition, pregnant women have frequent contacts with antenatal clinic staff during their pregnancy, therefore increasing the chance that their smoking status will be identified and reported in their birth record.

Although results from the logistic regression analyses suggest that the odds ratios were independent of the source of the smoking information, this research supports the use of record linkage to define maternal smoking in analyses of routinely collected data. This study has shown that linked data differentially increased the ascertainment of smokers in not only women who are at higher risk of adverse outcomes (older age, multi-fetus pregnancy, diabetes, hypertension) but also in women who are known to under-report their smoking status (older age, married, immigrants from non-English speaking countries, and high socio-economic status).[23,26,27] Factors contributing to voluntary disclosure of smoking among pregnant women may include contextual circumstances in which women describe their smoking status, clarity and wording of questions being asked about smoking, and motivations of women as mediated by social desirability, cultural and social norms.[51] In addition, it is also important that linking the records has resulted in a greater detection of smokers among women who gave birth in private and rural hospitals, given existing evidence that the data recording in these hospitals is less likely to be complete than in public or urban hospitals.[17,52]

Comparing the two enhanced algorithms, it was found that additional information from antenatal admissions provided little enhanced benefit. This is in contrast to a report from another study that the use of antenatal admission records significantly increased the identification of pre-existing chronic diagnoses, by amounts ranging from 6.0% for diabetes, 17.9% for thyroid disorders, 30.5%

for psychiatric disorders, and 44.2% for asthma or chronic obstructive pulmonary disease, to 223.9% for chronic renal disease.[53] The disparity between the two studies is likely to be due to the difference in nature of conditions under investigation. Chronic conditions often complicate the pregnancy, and need to be managed in hospital, and, as a result, present greater opportunities for identification during antenatal admissions. Meanwhile, smoking is of concern only when it interferes with patient management.[37]

Although no separate data source was available to compare the enhanced identification of smokers, by benchmarking findings against the population-based surveys this research has demonstrated that the enhanced algorithms produce estimates of maternal smoking prevalence that are similar to the expected rates of smoking in women of reproductive age. Low false positives for smoking in birth records and hospital records[24,25,54] suggest that the use of linked data reduced the number of smokers who were incorrectly reported as non-smokers, rather than incorrectly identifying non-smokers as smokers.

CONCLUSION

Combining data from routinely collected birth records and records of hospital admissions for the birth delivery significantly enhanced the identification of women who smoked during pregnancy, thus generating better estimates of maternal smoking prevalence. Greatest enhancement was among older mothers, immigrants from non-English speaking countries, those who lived in high socio-economic areas, had a high-risk pregnancy, and gave birth in private or rural hospitals. Studies using a single source of routinely collected data should consider making adjustments for the differential misclassification of smoking status in certain obstetric populations.

Acknowledgements

We thank the NSW Department of Health for access to the population health data and the NSW Centre for Health Record Linkage for linking the data sets. This work was conducted as part of The Smoking MUMS study, which is funded by a National Health and Medical Research Council Project Grant (#1028543). CR is supported by a National Health and Medical Research Council Senior Research Fellowship (#1021025). AH is supported by a National Health and Medical Research Council Early Career Fellowship (#1013287).

Competing interests

The authors declare that they have no competing interests.

REFERENCES

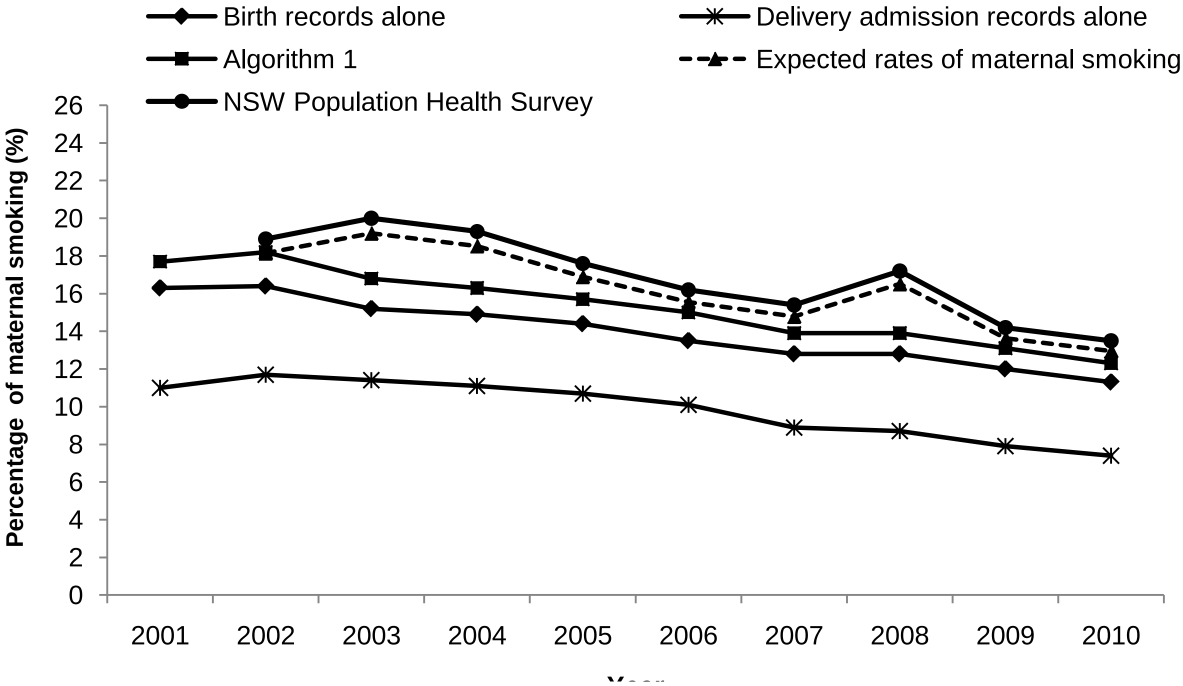
1. Cnattingius S. The epidemiology of smoking during pregnancy: smoking prevalence, maternal characteristics, and pregnancy outcomes. Nicotine Tob Res. 2004;6 Suppl 2:S125-40.
2. Australian Health Ministers' Advisory Council. Clinical Practice Guidelines: Antenatal Care - Module 1. Canberra: Australian Government Department of Health and Ageing 2012.
3. Mohsin M, Jalaludin B. Influence of previous pregnancy outcomes and continued smoking on subsequent pregnancy outcomes: An exploratory study in Australia. BJOG. 2008;115(11):1428-35.
4. Lumley J, Chamberlain C, Dowswell T, Oliver S, Oakley L, Watson L. Interventions for promoting smoking cessation during pregnancy. Cochrane Database Syst Rev. 2009(3):CD001055.
5. Schneider S, Huy C, Schutz J, Diehl K. Smoking cessation during pregnancy: a systematic literature review. Drug Alcohol Rev. 2010;29(1):81-90.
6. Dietz PM, Adams MM, Rochat RW, Mathis MP. Prenatal smoking in two consecutive pregnancies: Georgia, 1989-1992. Matern Child Health J. 1997;1(1):43-51.
7. Hoff GL, Cai J, Okah FA, Dew PC. Changes in smoking behavior between first and second pregnancies. Am J Health Behav. 2007;31(6):583-90.
8. Cnattingius S, Akre O, Lambe M, Ockene J, Granath F. Will an adverse pregnancy outcome influence the risk of continued smoking in the next pregnancy? Am J Obstet Gynecol. 2006;195(6):1680-6.
9. Hauge LJ, Aaro LE, Torgersen L, Vollrath ME. Smoking during consecutive pregnancies among primiparous women in the population-based Norwegian Mother and Child Cohort Study. Nicotine Tob Res. 2013;15(2):428-34.
10. Egebjerg Jensen K, Jensen A, Nohr B, Kruger Kjaer S. Do pregnant women still smoke? A study of smoking patterns among 261,029 primiparous women in Denmark 1997-2005. Acta Obstet Gynecol Scand. 2008;87(7):760-7.
11. Mohsin M, Bauman AE, Forero R. Socioeconomic correlates and trends in smoking in pregnancy in New South Wales, Australia. J Epidemiol Community Health. 2011;65(8):727-32.

12. Centers for Disease Control and Prevention. Tobacco use and pregnancy 2013 [cited 2013 8 March]. Available from: http://www.cdc.gov/reproductivehealth/tobaccousepregnancy/.
13. Al-Sahab B, Saqib M, Hauser G, Tamim H. Prevalence of smoking during pregnancy and associated risk factors among Canadian women: a national survey. BMC Pregnancy Childbirth. 2010;10:24.
14. Scollo M, Winstanley M. Tobacco in Australia: Facts and issues (4th edition) Cancer Council Victoria; 2012 [cited 2013 May 10]. Available from: www.TobaccoInAustralia.org.au
15. Zwar N, Richmond R, Borland R, Peters M, Litt J, Bell J, et al. Supporting smoking cessation: a guide for health professionals. Melbourne: The Royal Australian College of General Practitioners, 2011.
16. The Royal Australian and New Zealand College of Obstetricians and Gynaecologists. Women and Smoking (College Statement C-Gen 5). Melbourne 2011.
17. Sutherland JM, Steinum O. Hospital factors associated with clinical data quality. Health Policy. 2009;91(3):321-6.
18. Kelman CW, Bass AJ, Holman CDJ. Research use of linked health data: A best practice protocol. Aust N Z J Public Health. 2002;26(3):251-5.
19. Havard A, Jorm LR, Preen D, Daube M, Kemp A, Einarsdóttir K, et al. The Smoking MUMS (Maternal Use of Medications and Safety) Study: protocol for a population-based cohort study using linked administrative data. BMJ Open. 2013;3(9).
20. Centre for Epidemiology and Research. New South Wales Mothers and Babies 2009. Sydney: NSW Ministry of Health 2011.
21. Li Z, Zeki R, Hilder L, Sullivan E. Australia's mothers and babies 2010. Canberra: AIHW National Perinatal Epidemiology and Statistics Unit 2012.
22. Dietz PM, Adams MM, Kendrick JS, Mathis MP. Completeness of ascertainment of prenatal smoking using birth certificates and confidential questionnaires: variations by maternal attributes and infant birth weight. PRAMS Working Group. Pregnancy Risk Assessment Monitoring System. Am J Epidemiol. 1998;148(11):1048-54.
23. Land TG, Landau AS, Manning SE, Purtill JK, Pickett K, Wakschlag L, et al. Who underreports smoking on birth records: A Monte Carlo predictive model with validation PLOS ONE. 2012;7(4):e34853.
24. Taylor L, Pym M, Bajuk B, Sutton L, Travis S, Banks C. Validation study NSW Midwives Data Collection 1998. Sydney: NSW Health Department 2000.
25. Taylor LK, Travis S, Pym M, Olive E, Henderson-Smart DJ. How useful are hospital morbidity data for monitoring conditions occurring in the perinatal period? Aust N Z J Obstet Gynaecol. 2005;45(1):36-41.
26. Allen AM, Dietz PM, Tong VT, England L, Prince CB. Prenatal smoking prevalence ascertained from two population-based data sources: birth certificates and PRAMS questionnaires, 2004. Public Health Rep. 2008;123(5):586-92.
27. Shipton D, Tappin DM, Vadiveloo T, Crossley JA, Aitken DA, Chalmers J. Reliability of self reported smoking status by pregnant women for estimating smoking prevalence: a retrospective, cross sectional study. BMJ. 2009;339:b4347.
28. Dietz PM, Homa D, England LJ, Burley K, Tong VT, Dube SR, et al. Estimates of nondisclosure of cigarette smoking among pregnant and nonpregnant women of reproductive age in the United States. Am J Epidemiol. 2011;173(3):355-9.
29. Mohsin M, Forero R, Berthelsen A, Bauman A, Jalaludin B, Gyaneshwar R. Social influences for smoking in pregnancy in south western Sydney antenatal clinic attendees. Aust N Z J Obstet Gynaecol. 2007;47(3):207-12.

30. Lindqvist R, Lendahls L, Tollbom O, ABerg H, Hakansson A. Smoking during pregnancy: comparison of self-reports and cotinine levels in 496 women. Acta Obstet Gynecol Scand. 2002;81(3):240-4.
31. Ford RP, Tappin DM, Schluter PJ, Wild CJ. Smoking during pregnancy: how reliable are maternal self reports in New Zealand? J Epidemiol Community Health. 1997;51(3):246-51.
32. Roberts CL, Cameron CA, Bell JC, Algert CS, Morris JM, Roberts CL, et al. Measuring maternal morbidity in routinely collected health data: development and validation of a maternal morbidity outcome indicator. Med Care. 2008;46(8):786-94.
33. Roberts CL, Bell JC, Ford JB, Hadfield RM, Algert CS, Morris JM. The accuracy of reporting of the hypertensive disorders of pregnancy in population health data. Hypertension in pregnancy : official journal of the International Society for the Study of Hypertension in Pregnancy. 2008;27(3):285-97.
34. Lydon-Rochelle MT, Holt VL, Nelson JC, Cardenas V, Gardella C, Easterling TR, et al. Accuracy of reporting maternal in-hospital diagnoses and intrapartum procedures in Washington State linked birth records. Paediatr Perinat Epidemiol. 2005;19(6):460-71.
35. Lydon-Rochelle MT, Holt VL, Cardenas V, Nelson JC, Easterling TR, Gardella C, et al. The reporting of pre-existing maternal medical conditions and complications of pregnancy on birth certificates and in hospital discharge data. Am J Obstet Gynecol. 2005;193(1):125-34.
36. Salihu HM, Wilson RE. Epidemiology of prenatal smoking and perinatal outcomes. Early Hum Dev. 2007;83(11):713-20.
37. National Centre for Classification in Health. The International Statistical Classification of Diseases and Related Health Problems, Tenth Revision, Australian Modification (ICD-10-AM)-Fifth Edition-Tabular list of diseases and Alphabetic index of diseases. Sydney: National Centre for Classification in Health, Faculty of Health Sciences, The University of Sydney; 2006.
38. Centre for Health Record Linkage. Master Linkage Key [Online]. Centre for Health Record Linkage; 2011 [cited 2012 April 1]. Available from: http://www.cherel.org.au/our-services.
39. Bentley JP, Ford JB, Taylor LK, Irvine KA, Roberts CL. Investigating linkage rates among probabilistically linked birth and hospitalization records. BMC Med Res Methodol. 2012;12:149.
40. Australian Bureau of Statistics. Standard Australian Classification of Countries (SACC) Australia, Second edition. Canberra: ABS 2008. Report No.: ABS Cat. No. 1269.0.
41. Australian Bureau of Statistics. 2033.0.55.001. Census of population and housing: Socio-Economic Indexes for Areas (SEIFA), Australia 2006 [Online]. ABS; 2008 [cited 2011 March 13]. Available from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/2033.0.55.001Main+Features12006?OpenDocument.
42. Aliyu M, Lynch ON, Wilson R, Alio A, Kristensen S, Marty P, et al. Association between tobacco use in pregnancy and placenta-associated syndromes: a population-based study. Arch Gynecol Obstet. 2011;283(4):729-34.
43. Chan A, Keane RJ, Robinson JS. The contribution of maternal smoking to preterm birth, small for gestational age and low birthweight among Aboriginal and non-Aboriginal births in South Australia. Med J Aust. 2001;174(8):389-93.
44. Salihu HM, Shumpert MN, Aliyu MH, Kirby RS, Alexander GR. Smoking-associated fetal morbidity among older gravidas: a population study. Acta Obstet Gynecol Scand. 2005;84(4):329-34.
45. Salihu HM, Sharma PP, Getahun D, Hedayatzadeh M, Peters S, Kirby RS, et al. Prenatal tobacco use and risk of stillbirth: A case-control and bidirectional case-crossover study. Nicotine Tob Res. 2008;10(1):159-66.

46. Ananth CV, Smulian JC, Demissie K, Vintzileos AM, Knuppel RA. Placental abruption among singleton and twin births in the United States: Risk factor profiles. Am J Epidemiol. 2001;153(8):771-8.
47. van den Berg G, van Eijsden M, Galindo-Garre F, Vrijkotte TGM, Gemke RJBJ. Smoking overrules many other risk factors for small for gestational age birth in less educated mothers. Early Hum Dev. 2013;89(7):497-501.
48. Raymond EG, Cnattingius S, Kiely JL. Effects of maternal age, parity, and smoking on the risk of stillbirth. Br J Obstet Gynaecol. 1994;101(4):301-6.
49. New South Wales Department of Health. New South Wales Population Health Surveys 2012 [cited 2013 May 21]. Available from: http://www0.health.nsw.gov.au/publichealth/surveys/phs.asp.
50. Mohsin M, Bauman AE. Socio-demographic factors associated with smoking and smoking cessation among 426,344 pregnant women in New South Wales, Australia. BMC public health. 2005;5:138.
51. Shaffer HJ, Eber GB, Hall MN, Bilt JV. Smoking behavior among casino employees: Self-report validation using plasma cotinine. Addict Behav. 2000;25(5):693-704.
52. Tran DT, Jorm L, Lujic S, Bambrick H, Johnson M. Country of birth recording in Australian hospital morbidity data: accuracy and predictors. Aust N Z J Public Health. 2012;36(4):310-6.
53. Chen JS, Roberts C, Simpson J, Ford J. Use of hospitalisation history (lookback) to determine prevalence of chronic diseases: Impact on modelling of risk factors for haemorrhage in pregnancy. BMC Med Res Methodol. 2011;11(1):68. DOI:10.1186/471-2288-11-68.
54. Henderson T, Shepheard J, Sundararajan V. Quality of diagnosis and procedure coding in ICD-10 administrative data. Med Care. 2006;44(11):1011-9.

Figure 1: Prevalence of maternal smoking according to enhanced and unenhanced identification of smokers, and rates of smoking in the general female population in NSW, 2001-2010



- NSW Population Health Survey: Age-adjusted rates of daily or occasional smoking among females (≥16 years). Rates and number of respondents: 2002 (18.9%, 7388), 2003 (20.0%, 7759), 2004 (19.3%, 5855), 2005 (17.6%, 6962), 2006 (16.2%, 4802), 2007 (15.4%, 8178), 2008 (17.2%, 6268), 2009 (14.2%, 6696) and 2010 (13.5%, 6475).
- In 2001, PHS was not conducted. Birth records and hospital data were available for the last 6 months.
- Prevalence estimates by algorithm 1 and algorithm 2 are almost identical.

Table 1: Number of pregnant smokers identified by the enhanced algorithms (N) and percentage increase compared to unenhanced identification (%)

		Smokers identified by algorithms and percentage increase				
		Algorithm 1*			Algorithm 2**	
	Number of pregnancies	N	Increase over birth records alone (%)	Increase over delivery admissions alone (%)	N	Increase over algorithm 1 (%)
Overall	**846,039**	**127,612**	**9.6**	**54.6**	**130,020**	**1.9**
Year of delivery						
2001[§]	41,955	7,426	8.4	60.8	7,536	1.5
2002	83,923	15,241	10.8	54.6	15,446	1.3
2003	84,439	14,195	10.7	48.0	14,440	1.7
2004	83,070	13,573	9.9	46.6	13,829	1.9
2005	87,724	13,776	9.4	47.2	14,080	2.2
2006	89,862	13,455	10.8	48.9	13,722	2.0
2007	93,604	13,038	8.8	57.1	13,313	2.1
2008	93,682	13,040	9.0	59.9	13,313	2.1
2009	94,148	12,317	8.6	65.8	12,563	2.0
2010	93,632	11,551	9.2	65.9	11,778	2.0
Maternal age[†]						
Under 20 years	32,058	13,221	6.6	50.3	13,495	2.1
20-24 years	119,830	33,412	7.8	53.5	34,041	1.9
25-29 years	232,811	34,874	10.1	54.6	35,496	1.8
30-34 years	279,292	28,003	11.5	56.7	28,517	1.8
≥35 years	182,032	18,098	11.7	56.4	18,467	2.0
Country of birth						
English-speaking	649,954	118,747	9.1	52.6	120,909	1.8
Asian	106,572	2,293	21.1	115.7	2,348	2.4
Other	89,513	6,572	17.0	78.2	6,763	2.9
Marital status						
Married, de-facto	697,659	69,530	12.1	58.5	71,021	2.1
Never married	130,336	52,601	6.6	50.0	53,415	1.5
Widow, separated, others	18,044	5,481	8.3	51.0	5,584	1.9
Socioeconomic status (SEIFA)[†]						
1st quintile (highest)	193,586	10,014	18.8	77.2	10,339	3.2
2nd quintile	169,335	20,521	12.4	53.5	20,957	2.1
3rd quintile	190,504	33,963	9.3	49.7	34,567	1.8
4th quintile	146,294	32,222	7.4	52.6	32,739	1.6
5th quintile (lowest)	137,654	29,295	8.1	57.7	29,806	1.7
Maternity hospital[†]						
Private	203,197	5,472	28.6	96.3	5,716	4.5
Tertiary	231,773	32,708	14.9	53.4	33,584	2.7
Regional urban	148,088	32,359	8.7	46.3	33,018	2.0
District urban	244,294	52,247	6.1	55.0	52,825	1.1
Rural	18,675	4,819	3.6	83.7	4,870	1.1

Parity†						
Primipara	353,943	44,799	10.7	66.2	45,868	2.4
One	283,672	35,819	10.8	51.1	36,449	1.8
Two	128,903	23,214	8.9	48.9	23,616	1.7
Three or more	78,308	23,620	6.6	45.8	23,923	1.3
Plurality						
Singleton	832,859	125,813	9.6	54.6	128,140	1.8
Twins	12,933	1,770	14.1	53.9	1,849	4.5
Triplets/ quadruplets	247	29	16.0	61.1	31	6.9
Diabetes* **						
Yes	44,218	4,982	14.3	52.2	5,128	2.9
No	801,821	122,630	9.5	54.7	124,892	1.8
Hypertensive disorders* **						
Yes	78,169	8,762	14.1	52.1	9,062	3.4
No	767,870	118,850	9.3	54.7	120,958	1.8
Preterm birth†						
Yes (<37 weeks)***	55,708	12,635	9.0	50.7	13,028	3.1
No	790,323	114,976	9.7	55.0	116,991	1.8
Stillbirth†						
Stillbirth	5,038	1,122	8.8	63.8	1,151	2.6
Live birth	840,599	126,390	9.6	54.5	128,767	1.9
Placental abruption						
Yes	3,772	1,048	10.5	44.0	1,072	2.3
No	842,267	126,564	9.6	54.7	128,948	1.9
Birthweight for gestational age‡						
Appropriate	669,016	96,753	9.9	55.8	98,592	1.9
Small	78,997	21,678	7.0	47.0	21,967	1.3
Large	91,550	7,814	14.1	60.8	8,059	3.1

§ Data for 2001 were available for the last six months

*: Smoker as indicated by birth record or delivery admission record

**: Smoker as indicated by birth record, delivery admission record, or antenatal admission record

***: Diabetes included gestational and pre-existing diabetes. Hypertensive disorders included chronic hypertensive and pregnancy-induced hypertensive diagnoses. Preterm birth included medically indicated and spontaneous delivery <37 weeks.

†: Total number of pregnancy less than 846,039 due to missing data.

‡: Among live births, total number of pregnancy less than 840,599 due to missing data.
Small birthweight: <10th percentile, Large: >90th percentile.
SEIFA: Socio-Economic Indexes for Area, quintile scores of the Index of Relative Socio-economic Disadvantage

Table 2: Odds ratios of pregnancy outcomes according to identification methods in this study in comparison with published results from other population-based studies.

	Sample size*; data source, study period	Estimates and 95% confidence intervals			
		Placental abruption	Preterm birth	Stillbirth	Small for gestational age
This study	832,859; NSW birth records and hospital admission records, 2001-2010				
Birth records alone, *OR***		1.94 (1.78-2.12)	1.67 (1.63-1.72)	1.38 (1.27-1.49)	2.46 (2.41-2.52)†
Delivery admission records alone, *OR***		1.97 (1.79-2.16)	1.59 (1.55-1.64)	1.21 (1.10-1.32)	2.28 (2.23-2.33)†
Algorithm 1, *OR***		2.04 (1.87-2.22)	1.68 (1.64-1.73)	1.38 (1.28-1.49)	2.44 (2.39-2.49)†
Algorithm 2, *OR***		2.07 (1.90-2.25)	1.73 (1.68-1.77)	1.40 (1.30-1.52)	2.43 (2.38-2.48)†
Previous population-based studies					
Cnattingius 2004,[1] RR or OR	A review of literature	1.40 to 2.40	1.20 to 1.60	1.20 to 1.80	1.50 to 2.90
Aliyu 2011,[42] *OR*	312,505; US Missouri birth certificates, 1989-2005	1.67 (1.59-1.74)	1.19 (1.17-1.21)	1.20 (1.13-1.28)	2.27 (0.24-2.30)
Chan 2001,[43] *RR*	36,059; South Australia Perinatal Data Collection, 1998-1999		1.64 (1.51-1.80)		2.28 (2.14-2.43)
Salihu 2005,[44] *OR*	7,792,990; US Perinatal data, 1995-1997		1.30 (1.20-1.30)		2.30 (2.20-2.40)
Salihu 2008,[45] *OR*	1,444,378; US Missouri birth certificates, 1987-1997	1.34 (1.26-1.43)			
Ananth 2001,[46] *RR*	7,465,858; US National birth and death certificates, 1995-1996	1.97 (1.92-2.02)			
van den Berg 2013,[47] OR	3,793; Netherlands population-based cohort study, 2003-2004				3.06 (2.11–4.43)
Raymond 1994,[48] *OR*	638,242; Sweden Medical Birth Register, 1983-1989			1.40 (1.20-1.40)	

OR: Odds ratios, RR: Relative risks *: Among singleton pregnancies **: Adjusted for year of birth, maternal age, country of birth, marital status, private health insurance, residential remoteness, SEIFA, parity, diabetes, and hypertensive disorders. †: Among live births