

A Thesis submitted in partial fulfilment of the requirements for the degree of

Master of Science

Department of Epidemiology

The Faculty of Graduate Studies

University of Ottawa

Title

**“VALIDITY OF ADMINISTRATIVE DATABASE FOR REPORTING
PRE-ECLAMPSIA”**

By

SVETLANA SHACHKINA
MD (Honours), PhD (Medicine)
Student No [REDACTED]

Ottawa, Ontario

May 2012

Abstract

Background: Pre-eclampsia (PET) is one of the major causes of maternal and neonatal morbidity and mortality¹. Misclassification of PET can lead to biased or erroneous results in epidemiologic studies resulting in false conclusions.

Objectives: The objectives of this thesis are to determine the validity of PET diagnosis in pregnant women in administrative database using the ICD-10-CA codes, to explore the nature of misclassification, and to estimate whether misclassification of PET diagnosis in administrative database may result in biased conclusions.

Methods: Pregnant women who participated in the Ottawa and Kingston (OaK) Birth Cohort study and delivered in the Ottawa Hospital were included in the study. All cases with hypertensive disorder of pregnancy in the study population were adjudicated to confirm diagnosis of PET. This adjudicated dataset was used as a reference standard. The PET incidence in hospital discharge database was compared with PET incidence calculated from the reference standard database.

Results: 2887 of the requested charts were available for review. The PET incidence was much lower in administrative database (1.47%) than in the OaK Birth Cohort Study (3.6%). The results of the study demonstrated that hospital discharge database via ICD-10-CA was not very sensitive to determine incidence of PET since sensitivity of ICD-10-CA diagnostic codes for PET was low (35.92% with 95% Confidence Intervals (CI): 26.7; 45.9) but specificity, PPV, and NPV were high. The majority of misclassified cases belonged to the category (according to the proposed classification) “PET pregnancies coded with incorrect ICD-10-CA code” (78.88%) followed by the category “Pregnancies affected by PET coded as normal” (14.08%).

Conclusion: Using hospital discharge database and ICD-10-CA coding to determine incidence of PET in certain settings may yield low sensitivity. Researchers should validate the results when using the hospital discharge database for PET research to ensure that the findings based on analyses of such data demonstrate what they claimed to demonstrate.

Table of Contents

Abstract	1
Table of Contents	3
Acknowledgements	5
Dedication	6
List of Tables.....	7
List of Figures	9
List of Abbreviations.....	10
Chapter 1 General introduction	12
Chapter 2 Background.....	16
Chapter 3 Proposed Approach.....	34
Chapter 4 Project 1	35
Inter-rater Agreement Estimation.....	35
4.1 Aim of the Project	35
4.2 Hypothesis	35
4.3 Background	35
4.4 Methods	43
4.5 Adjudication Results	50
4.6 Discussion	50
4.7 Conclusion.....	51
Chapter 5 Project 2	52
Validity of PET diagnosis in an administrative database	52
5.1 Aim.....	52
5.2 Hypothesis	52
5.3 Background	52
5.4 Methods	58
5.5 Results	67
5.6 Discussion	83
5.7 Conclusion.....	93
Chapter 6 Project 3	94
Evaluation of the Effect Outcome Misclassification on the Results	94
6.1 Aim.....	94

6.2 Hypothesis	94
6.3 Background	94
6.4 Methods	98
6.5 Results	102
6.6 Discussion	110
6.7 Conclusion.....	112
Chapter 7 General discussion	112
Chapter 8 Significance of the project	117
References	119
Appendices	134

Acknowledgements

I would like to thank my supervisors, Dr. Mark Walker and Dr. Shi Wu Wen – for mentorship, guidance and encouragement throughout this process.

I would like to acknowledge the useful contributions from Dr. Mark Rodger.

I would also like to thank all staff of Ottawa Hospital Health Record for helping me with chart pulls and special thanks to Spencer Holmes and Celine Groulx at the General Campus and Lilly Polesello at the Civic campus who cheerfully pulled charts and responded to my questions. Thanks to Health Records Analysts Sylvie Demers, Christine Leclair and Sheila Lessard for their eagerness to help in any possible way.

Many thanks to my colleagues who encouraged me with my studies.

I owe so much to my friends Tom Eades, Francois Tshibemba, Deshayne Fell and Liu Ning who provided me with review, suggestions, and encouragement. I would also like to thank my husband for his patience and support during my studies.

Dedication

This thesis is dedicated to my family, and, in particular, to my parents.

They always believed in me no matter what and always encouraged me to go further in seeking knowledge. Thank you and I love you.

List of Tables

Table 2.1 Evolution of the PET definition.....	28
Table 3.2 SOGC grouping of HDP and their approximation in the ICD-10-CA.....	33
Table 4.1 Proportions joined judgments of two judges on a scale with r categories.....	40
Table 4.2 Level of agreement and Kappa value.....	42
Table 5.1 Two-by-two table used for the Ottawa hospital database accuracy calculations.....	64
Table 5.2 Hypertensive diseases of pregnancy frequencies according to the reference standard data.....	69
Table 5.3 Hypertensive diseases of pregnancy frequencies according to the Ottawa hospital discharge database.....	69
Table 5.4 Two-by-two table of PET from the reference standard data versus data from the Ottawa hospital discharge database	70
Table 5.5 Two-by-two table of HDP (PET and GH combined) from the reference standard data versus data from the Ottawa hospital discharge database	71
Table 5.6 Demographic and pregnancy history characteristics for the correctly classified/misclassified HDP cases.....	75
Table 5.7 Demographic and pregnancy history characteristics for the correctly classified/misclassified PET cases	76
Table 5.8 Incidence of PET/HDP based on parity in the reference standard data.....	77

Table 5.9 Incidence of PET/HDP based on parity in the Ottawa hospital discharge database.....	77
Table 5.10 Adjusted versus crude odds ratios related to predictors of PET misclassification.....	79
Table 5.11 Frequencies of different kinds of errors in the Ottawa hospital discharge database.....	83
Table 6.1 Socio-demographic characteristics of women according to PET status..	104-105
Table 6.2 Adjusted versus crude odds ratios related to predictors of PET.....	106-107
Table 6.3 Adjusted ORs of association between PET risk factors and PET incidence according to the reference standard data versus data obtained through the Ottawa hospital discharge database.....	109
Table A3.1 Agreement calculations for 2 categories and 2 raters.....	145
Table A3.2 Agreement calculations for “K” categories for 2 raters.....	146
Table A3.3 Agreement calculations for 3 categories and 3 raters.....	149-150
Table A3.4 Agreement calculations for 3 categories for 3 raters (without the subjects to whom at least one rater did not assign a category)	151-152
Table A4.1 Number of cases (or controls) for expected sensitivities (or specificities).....	154

List of Figures

Figure 4.1 Adjudication Process.....	50
Figure 5.1 Causes of coding errors.....	66

List of Abbreviations

ACOG	American Congress of Obstetricians and Gynecologists
ALT	Alanine Aminotransferase
AST	Aspartate Aminotransferase
BMI	Body Mass Index (kg/m ²)
BP	Blood Pressure
CDC	Centre for Disease Control and Prevention
CI	Confidence Interval
CIHI	Canadian Institute for Health Information
CRF	Case Report Form
CS	Caesarean Section
DAD	Discharge Abstract Database
DRG	Drug-Related Groups
EDD	Estimated Date of Delivery
g/day	gram per day
GH	Gestational hypertension
ICD	International Classification of Disease
ICD-9-CM	International Classification of Diseases, ninth edition, Clinical Modification
ICD-10-CA	International Classification of Diseases, tenth edition, Canadian
IUGR	Intrauterine Growth Restriction
HDP	Hypertensive Disease of Pregnancy
HELLP	Hemolysis, Elevated Liver Enzymes, Low Platelets

LDH	Lactate Dehydrogenase
LOS	Length of Stay
LR:	Likelihood ratio
mmHg	millimetre of mercury
MTHFR	5, 10-methylenetetrahydrofolate reductase
NPV	Negative Predictive Value
OaK	Ottawa and Kingston
OR	Odds ratio
PET	Preeclampsia
PIH	Pregnancy Induced Hypertension
PA	Placenta Abruptio
PPV	Positive Predictive Value
RCT	Randomized Controlled Trial
SOGC	Society of Obstetricians and Gynecologists of Canada
UK	United Kingdom
USA	United States of America
VTE	Venous Thromboembolism
WHO	World Health Organization

Chapter 1 General introduction

Pre-eclampsia (PET) is one of the major causes of maternal and neonatal morbidity and mortality. [1] Reported rates of PET vary between 3.0 to 10%, depending on the studied population and the clinical definitions of PET. [2, 3] According to a WHO report on the global burden of hypertensive disorders in pregnancy in year 2000, the incidence of PET ranged from 0.8% for developed countries to 3.4% for developing countries. [4]

The etiology of PET is not completely understood, but it appears that the placenta may play a central role in the pathogenesis of PET. [5, 6] The reported risk factors for PET include nulliparity, family history of or past history of PET, pre-existing diabetes, increased pre-pregnancy body mass index (BMI), multiple pregnancy, extreme maternal age (<20 or >40 years), renal disease, chronic hypertension and chronic autoimmune disease. [7]

PET can be a devastating and life threatening condition for both mother and the baby. It can cause problems in the mother's liver, kidneys, and brain, and abnormalities of the clotting system. Rare but serious complications include the following: eclampsia (the occurrence of seizures), stroke, HELLP syndrome (hemolysis, elevated liver enzymes, and low platelets), and disseminated intravascular coagulation. These complications are associated with an increased risk of maternal death. There are also increased risks for the fetus and newborn associated with this condition. The most common abnormalities are poor growth due to inadequate blood supply through the placenta that is not functioning properly, and the problems of prematurity mostly associated with delivery before the term (preterm delivery) in order to protect health of the mother or the fetus. [8, 9]

Complications associated with premature birth include respiratory distress, apnea, jaundice, kernicterus, feeding difficulties, hypoglycemia, seizures, periventricular leucomalacia, and prolonged hospitalization. [10] These infants require modern neonatal facilities and advanced care, both of which are expensive and not easily accessible in some countries. [9]

Perinatal mortality is high after PET, and especially after complications of PET. It is especially relevant for developing countries where one-quarter of stillbirths and neonatal deaths are associated with PET. Even though perinatal mortality is lower in developed countries, it still exists; for example, in the United Kingdom overall 6% of children born to women who had eclampsia die. Overall, for women with PET, mortality for the infants is three times higher in developing countries than in developed countries. [9]

PET/eclampsia is also associated with substantial health problems later in life for both women and their children. There is growing evidence that women who have had PET during their pregnancy are at increased risk of cardiovascular diseases later in life. [11, 12] Children of mothers who developed PET during pregnancy are at an increased risk of cerebral palsy compared to the children of mothers who did not suffer from this condition during pregnancy. [9]

While the global financial burden is difficult to calculate, the immediate and long-term annual costs have been estimated in billions of dollars. [9]

Despite the widespread recognition of the importance of PET as a cause of morbidity and mortality worldwide, we continue to have only sketchy information on the true incidence of PET, because most reports document PET occurrence using different definitions. [13]

The diagnosis of PET is difficult due to variations in standards, confusing nomenclature and definitions. The nomenclature of hypertension during pregnancy is pregnancy-induced hypertension (PIH) or gestational hypertension (GH), transient hypertension of pregnancy, pre-eclampsia, and superimposed pre-eclampsia. [14, 15] But it is reported that many authors use the term PIH/GH interchangeably with PET. PIH/GH and PET share the same clinical symptom such of raised blood pressure (BP) in pregnancy but they represent two very different clinical entities with important monitoring and treatment differences. [16]

Villar et al. [17] estimated that overall, hypertension complicates approximately 5% of all pregnancies (approximately 11% of all first pregnancies). Of these, approximately half are due to or associated with PET. The authors identified that the heterogeneity of classification systems is one issue that prevents a summary of the epidemiological data and the comparison of the reports. In general, the definition proposed by Davey and MacGillivray in 1988 [18] remains the most commonly used to diagnose hypertension in pregnancy. More recently, varying definitions have been proposed by the Australasian Society [19], the National High Blood Pressure Education Programme (NHBPEP) Working Group [20], and the Canadian Hypertension Society [21]. The authors suggested that an effort should be made to arrive to a consensus on unified definition of PET. Only the components of definitions that have been shown to increase substantive risk should be retained. Misclassification of PET can lead to biased results in epidemiologic studies resulting in flawed conclusions. If the data used to derive incidence estimates are inaccurate and/or incomplete, the current impression of the magnitude of the health

problem posed by PET may be erroneous. An accurate incidence rate for PET is vital for improvement of outcomes for these women and their babies. [17]

There are different approaches to studying PET whether by designing large prospective studies or by making linkage to national administrative databases, such as hospital discharge databases where data are routinely collected for administrative purposes and coded in accordance with the International Classification of Disease (ICD). Administrative databases are appealing sources for clinical research because of low cost and easy access to a large number of records, thus administrative databases have been used increasingly in epidemiological and clinical outcome studies of hypertensive disorders of pregnancy. [22, 23] It is largely unknown how accurate these databases are because of very limited understanding of the validity of preeclampsia-related diagnoses collected in administrative databases. The importance to know how accurate they are in the estimation of PET incidence is paramount.

Chapter 2 Background

2.1 Overview of coding process

In Canada, health care providers document what happens to the patient while in hospital in the patient's medical chart. After the patient is discharged, the health record coders review all information in the medical chart to identify diagnoses and clinical procedures and summarize it into a discharge abstract using ICD codes. There are up to 25 diagnosis fields that identify the diagnosis most responsible for the admission, as well as the secondary diagnoses of the patient's condition. The ICD codes are abstracted to the administrative discharge abstract database (DAD). The DAD is a national database containing information related to hospital inpatient and day surgery events. In Canada this database is maintained by the Canadian Institute for Health Information (CIHI).

CIHI was established in 1994 by the federal, provincial and territorial ministers of health in response to the need to coordinate health information. CIHI's mandate is to provide accurate and timely information that is needed to establish sound health care policy and to effectively manage Canada's health care system. Currently, over four million records are submitted to the DAD annually. Records of inpatient stay submitted to the DAD represent 75% of all patient discharges in Canada. Each record in the DAD contains standard clinical, demographic and administrative data for the health services provided for each inpatient stay. [24, 25]

2.2 International Classification of Diseases

The development of the ICD has a history dating back to 1858 when the first International Statistical Congress requested an internationally applicable uniform nomenclature for causes of death. [27] The International Classification of Diseases is published by the World Health Organization (WHO) and used worldwide for morbidity and mortality statistics, reimbursement systems, and automated decision support in medicine. This system is designed to promote international comparability in the collection, processing, classification, and presentation of these statistics.

The tenth revision of the ICD system (ICD-10) was introduced in 1992 as an enhancement to ICD-9-CM. [27] The International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) which is an extension of the Ninth Revision, the International Classification of Diseases (ICD-9), is the official system used in the United States to classify and assign codes to health conditions and related information. [28] The new ICD-10 system has more codes and is more comprehensive than its predecessor ICD-9. [29, 30] Canada, Australia, Germany, and other countries have enhanced ICD-10 by adding more specific codes and have released country-specific ICD-10 versions. ICD-10-Canada (ICD-10-CA) is one of them; however, ICD-10-CA has maintained its comparability with ICD-10. The ICD-10-CA system has been implemented in an attempt to keep up with medical advancements and to establish national standards. [31] This coding system has been in place in Ontario since April 2002. [29]

2.3 The Importance of Accurate Coding

2.3.1 Introduction to Case Mix

Case Mix Group (CMG), is the foundation of CIHI's acute inpatient grouping, length of stay (LOS) and resource intensity weight methodologies. The patient's Most Responsible Diagnosis is used to assign the case to one of the 25 Major Clinical Categories. [25] Introduced in 1983, the Case Mix Groups system adapted the ICD-9-CM-based diagnosis-related groups system to accommodate ICD-9 classification systems. The result is groups of patients that are clinically similar and/or homogeneous with respect to hospital resources used. The creation of a Canadian grouper originated from the fact that health care management wanted:

- To improve the comparability of national health care data;
- To enhance the relationship between diagnoses, especially secondary diagnoses that contribute to longer hospital stay;
- To provide a tool for resource management based on Canadian health care data.

[25]

The CMG system is used for funding in other countries outside of Canada, inside Canada only being used in Ontario.

2.3.2 Purposes of Coded Data

Coded clinical data are held in databases which are subsequently accessed for a variety of purposes such as:

- A. *Funding*; the accuracy of clinical coding is crucial to the appropriate funding of health services when the funding mechanism is based on diagnosis-related group (DRG) which is equivalent of the CMG. [31] Marshall and Adema studied the

impact of accuracy of the coding on the funding of health care facility in Florida (USA). The authors demonstrated that the overall coding error in a health system in Florida was 84.5%, which led to a significant loss of funding. [32]

B. *Health statistics*; which influence social and medical research and health policy development and are used under the assumption that they are accurate. [31]

C. *Hospital discharge data* which are fundamental to planned national strategies designed to improve patient safety and quality of care.[31]

D. *Disease surveillance*; hospital discharge data are a critical resource for disease surveillance programs, which are used in turn for disease prevention and other public health programs. [31] For example, data used in the USA by the Center for Disease Control and Prevention (CDC) to compile its disease surveillance databases, are obtained directly from coded information which is collected on a regular basis from a vast number of sources in the USA. Accuracy and completeness of the coding recorded in these systems are essential for the planning of appropriate funding for many public health programs. [33] Schoenman et al. 2008 noted that the users of these databases are diverse, including government agencies, provider associations and individual providers, consumer organizations and patients, as well as, health care insurers and other purchasers. [34]

Coded hospital data are used for very important applications; therefore, the need for accuracy and reliability of hospital data coding is utmost importance. [31]

2.4 Validity of Diagnostic Information in Administrative Databases

Administrative databases are increasingly being used for research purposes to investigate health outcomes, disease prevalence/incidence, surveillance, in addition to other purposes. The primary advantages of using administrative data for these purposes are that these data are population-based, comprehensive, cost-efficient and free of biases associated with other data collection methods such as surveys (e.g., recall bias or non-response).[35] However, the accuracy of diagnosis in this data remains unclear and, therefore, limits the potential utilization of these databases for research.

The validity of the administrative data for different diagnoses has been evaluated by many researchers. Chiò et al. assessed the validity of the discharge of amyotrophic lateral sclerosis by using data from the Piemonte and Valle D'Aosta Register (Italy) as the gold standard. They found that hospital discharge records reflected the incidence of amyotrophic lateral sclerosis poorly and concluded that the data can be used only after clinical verification of the diagnosis. [36] Tairou et al. examined the validity of different sources of administrative data (death and stillbirth certificates and hospital discharge summaries) to identify neural tube defect cases in the population. They found that hospital discharge summaries had very good sensitivity but low positive predictive value and also had many coding errors due to imprecise diagnostic categories. They concluded that the combination of hospital discharge summaries and death and stillbirth certificates is a sensitive method for ascertaining of neural tube defect cases but must be complemented by review of medical records to exclude coding errors. [37] Dodds et al. used data from three administrative databases (Hospital Discharge Abstract Database, the Medical Services Insurance Physician Billing Database and the Mental Health Outpatient

information System Database) to estimate the prevalence of autism. They found that sensitivity based on autism spectrum disorder code in either database was 62.5% and concluded that administrative databases can potentially be a cost efficient source of data for autism surveillance, but additional data sources maybe needed to improve the sensitivity. [38] Quan et al. evaluated the validity of hospital discharge databases coded via ICD-10 in recording clinical conditions in four teaching hospitals in Alberta, Canada and compared it with the validity of hospital discharge databases coded via ICD-9. They found that data from hospital discharge databases coded by either ICD-10 or ICD-9 had similar validity in recording clinical condition information but there is no improvement in validity of hospital discharge databases coded by ICD-10 in comparison with ICD-9. [39] Humphries et al. evaluated whether accuracy of co-morbidity data obtained from administrative database (Canadian provincial hospitalization database) was comparable to the accuracy of information obtained through chart review which was considered a gold standard. They found that even though administrative database underestimated the prevalence of some co-morbid conditions, overall, the agreement between chart review and administrative data was good. [40] Ellekjær et al. assessed the accuracy of administrative database (hospital discharge database) in identifying cerebrovascular diseases in comparison with a stroke register in Norway which was considered as the gold standard in this study. They found that identifying cerebrovascular diseases by hospital discharge diagnoses (through ICD-9) led to a considerable overestimation of stroke in the target population. [41] Lofthus et al. examined the validity of electronic administrative database (hospital discharge database) of the hospitals in Oslo and the Norwegian Patient Register in identifying hip fracture cases using medical records as the

gold standard. They found that the hospital discharge database of one of the hospitals underestimated the number of hip fractures by 46%, while the other two overestimated the number of hip fractures by 17% and 19% respectively. For the national administrative database, an overall overestimation of 19% was found. They concluded that electronic administrative databases have questionable validity in identifying hip fracture and should be used cautiously in epidemiologic research. [42] Larsen et al. assessed the validity of diagnoses of venous thromboembolism (VTE) during pregnancy and postpartum, expressed as positive predictive value (PPV) in the Hospital Discharge Registry (administrative database) in North Jutland County of Denmark in comparison with chart review. The authors found that the overall PPV of VTE events in pregnancy was 79.3% (95% CI: 74.3–83.8). They concluded that even though overall PPVs of pregnancy-related VTE diagnoses were moderate to high, not all of the registered VTE events were related to pregnancy, thus, use of data from unvalidated administrative databases for epidemiological research may lead to biased results. [43] Rawson and D`Arcy warned that from a patient`s first contact with a physician to the recording of the encounter in an administrative database, there are numerous opportunities for errors that may affect the validity of the data. [44]

There are two main components of the validity in administrative data: the degree to which the information measures what it purports (measurement validity) and the extent to which results from a database can be generalized (external validity). [45] Rawson and D`Arcy tried to uncover the important issues in assessing the measurement validity of information in administrative databases. They recommended that a combined approach should be used when assessing measurement validity, consisting of comparisons of

administrative data with external sources (chart re-abstraction); or comparison with different data files (linkage), and, when appropriate, contextual consistency analyses (time-sequence relationships). These analyses should also be supported by careful consideration of the external validity of the data. In conclusion, they cautioned researchers to validate the results of each use of the administrative database to ensure that the findings based on the analyses of such data demonstrate what they claim to demonstrate. [46]

The results of many studies demonstrated that the validity of the data derived from administrative databases is questionable. Because the quality of such data is a concern, a group of researchers of health administrative data from Canada, the USA, Switzerland, Australia, China and UK came together to discuss and identify high-priority methodological research areas. The group discussion resulted in identification of thirteen potential areas of research to overcome challenges associated with using of administrative data in health services and population health research. [47]

2.5 Evolution of the Definition of Preeclampsia

Hypertensive disorders of pregnancy are subdivided into pre-existing hypertension, or chronic hypertension and pregnancy hypertension developing after 20 weeks of pregnancy. [1] Although there is an agreement on the classification of chronic hypertension, various classifications of pregnancy-induced hypertension exist.

Because knowledge about the pathophysiology of this condition is evolving, definitions of this disorder are evolving as well. For example, according to the *American Congress of Obstetricians and Gynecologists (ACOG) Technical Bulletin*, February 1986, #91:

"Hypertension is defined as a diastolic blood pressure of at least 90 mm Hg or a systolic blood pressure of at least 140 mm Hg, or a rise in the former of at least 15 mm Hg or in the latter of 30 mm Hg. These blood pressures must be manifested on at least two occasions 6 hours or more apart. PET is the development of hypertension with proteinuria (presence of 300 mg or more of protein in a 24-hour urine collection or a protein concentration of 1 g or more per litre in at least two random urine specimens collected 6 hours or more apart) or edema (a generalized accumulation of fluid of greater than 1+ pitting edema after 12 hours of bed rest or weight gain of 5 pounds or more in 1 week), or both, induced by pregnancy after the 20th week of gestation, and sometimes earlier, when there are extensive hydatidiform changes in the chorionic villi". [48] According to *Williams Obstetrics*, 1993 edition: "Preeclampsia is diagnosed by the development of hypertension plus proteinuria, or edema that is generalized and overt, or both. The diagnosis of pregnancy-induced hypertension is made when blood pressure is 140/90 mm Hg or greater." "Proteinuria is defined as 300 mg or more of urinary protein per 24 hours or 100 mg/dL or more in at least two random urine specimens collected 6 or more hours apart". [49] According to the *National High Blood Pressure Education Working Group Report on High Blood Pressure in Pregnancy*, "Preeclampsia is determined by increased blood pressure accompanied by proteinuria, edema, or both. Either of the following criteria suffice for the diagnosis of hypertension in this situation: (1) systolic blood pressure increases of 30 mm Hg or greater or (2) diastolic blood pressure increases of 15mm Hg or greater from early values (average of values before 20 weeks' gestation). If prior blood pressure is not known, readings of 140/90 mm Hg or greater after 20 weeks' gestation are considered sufficiently elevated to satisfy the blood pressure criteria of

preeclampsia. Note, however, that many young pregnant women will show the blood pressure increase required for the diagnosis of PET without increasing their pressure to 140/90 mm Hg". [20] "Proteinuria is defined as the excretion of 0.3 g or greater in a 24-hour specimen. This will usually correlate with 30 mg/dl ("1+ dipstick") or greater in a random urine determination. Proteinuria usually is a late sign in the course of preeclampsia; although it is nonspecific, its appearance greatly bolsters the diagnosis of preeclampsia. Edema is diagnosed as clinically evident swelling, but fluid retention may also be manifested as a rapid increase of weight without evident swelling". [18] According to the *Australasian Society for the Study of Hypertension 1993*, gestational hypertension was defined as the onset of hypertension (systolic BP $\geq$ 140 mmHg and/or diastolic BP (phase 4) $\geq$ 90 mmHg) after 20 weeks of gestation which returned to normal within 3 months of delivery. Hypertension in these women was confirmed after either overnight rest in hospital or following repeated BP measurements during a 4-5 hour day assessment unit visit. PET was diagnosed when one or more of the following maternal systemic complications accompanied hypertension:

- proteinuria $\geq$ 300 mg/day (or persistently $\geq$ 2+ on dipstick urinalysis if 24 hour urine was unavailable)
- renal impairment (plasma creatinine $\geq$ 100 μ mol/L)
- hepatic dysfunction (aspartate transaminase $\geq$ 50 IU/L and/or severe persistent epigastric pain)
- haematological abnormalities (haemolysis and/or platelet count $<$ 150 x 10⁹/L)
- cerebral disorder (visual scotomata, convulsions, hyperreflexia when accompanied by clonus) or

- severe hypertension (systolic BP ≥ 170 mmHg and/or diastolic BP ≥ 110 mmHg).

[17]

According to ACOG, 2002: criteria for the diagnosis of PET are blood pressure of 140 mm Hg systolic or higher or 90 mm Hg diastolic or higher that occurs after 20 weeks of gestation in a woman with previously normal blood pressure. Proteinuria is defined as urinary excretion of 0.3 g protein or higher in a 24-hour urine specimen. [14] The current definition according to *Canadian Clinical Practice Guideline*, 2008: in women with gestational hypertension, preeclampsia should be defined as new-onset proteinuria or one or more of the other adverse conditions. Hypertension in pregnancy should be defined as a diastolic BP of ≥ 90 mmHg, based on the average of at least two measurements, taken using the same arm. Women with a systolic BP of ≥ 140 mmHg should be followed closely for development of diastolic hypertension. Proteinuria should be strongly suspected when urinary dipstick proteinuria is $\geq 2+$. Proteinuria should be defined as ≥ 0.3 g/d in a 24-hour urine collection or ≥ 30 mg/mmol urinary creatinine in a spot (random) urine sample. [2]

It is evident that the definition of PET has evolved in the last 15 years (Table 2.1).

Table 2.1 Evolution of the PET definition

Criteria	ACOG Technical Bulletin, February 1986	Current definition
Hypertension	Hypertension is defined as a diastolic blood pressure of at least 90 mm Hg or a systolic blood pressure of at least 140 mm Hg, or a rise in the former of at least 15 mm Hg or in the latter of 30 mm Hg. These blood pressures must be manifested on at least two occasions 6 hours or more apart	Hypertension in pregnancy should be defined as a diastolic BP of ≥ 90 mmHg, based on the average of at least two measurements, taken using the same arm
Proteinuria	Presence of 300 mg or more of protein in a 24-hour urine collection or a protein concentration of 1 g or more per litre in at least two random urine specimens collected 6 hours or more apart	Proteinuria should be strongly suspected when urinary dipstick proteinuria is $\geq 2+$. Proteinuria should be defined as ≥ 0.3 g/d in a 24-hour urine collection or ≥ 30 mg/mmol urinary creatinine in a spot (random) urine sample
Edema	A generalized accumulation of fluid of greater than 1+ pitting edema after 12 hours of bed rest or weight gain of 5 pounds or more in 1 week	Excluded

2.6 PET Classifications Used in Current Research

Even though the definition of PET has evolved over the years, there are still various classification systems for PET used in research. Brown and Buddle showed in their research done on cohort of women with hypertension in pregnancy that depending on the classification of PET the percentage of women diagnosed with PET ranged from 16% to 77%. [50]

Chappel et al. conducted a systematic review of the literature to establish the criteria most commonly used by researchers to define PET and to compare the quality of description and accuracy of the definitions. They demonstrated that there was a lack of consistency in research papers in defining PET. A quarter of papers omit the definition of PET. Those authors who included the definition of PET in the papers often used different criteria for blood pressure measurement or proteinuria threshold. A number of papers referencing their choice of classification are mentioned. [52] Harlow and Brown also conducted a systematic review of 135 articles to establish features used by authors to define PET and its defining variables hypertension and proteinuria. The authors found that there was no consistency among researchers in defining either hypertension or proteinuria. The authors came to the conclusion that because scientific and research groups study different groups of preeclamptic women using different definitions for PET, the results derived from such studies cannot be considered truly comparable. [13] Brown and de Swiet acknowledged the importance of uniform classification for PET not only for clinicians but for researchers as well. It will allow comparing reported outcomes of this condition which will help in management of these patients. The authors recognized that unfortunately new-onset hypertension in pregnancy (in comparison with essential or chronic hypertension) is variably termed as “preeclampsia”, “toxaemia”, “gestational hypertension”, “transient hypertension”, “pregnancy-induced hypertension”, “oedema proteinuria hypertension gestosis”, or “hypertensive disorder of pregnancy”. According to Brown et al, “preeclampsia” and “pregnancy induced hypertension” are the most appropriate definitions. They also provide an example of the existence of different classifications used by different organizations such as the National High Blood Pressure

Education Program, the Australasian Society of Hypertension in Pregnancy and the American College of Obstetrician and Gynecologists. The authors proposed a new classification of hypertensive disorders of pregnancy which, while based on old classification, better reflects current understanding of pathophysiology of the disorder. [14] De Swiet pointed out that existing confusion among clinicians between PIH and PET could lead to misdiagnosis and erroneous coding of the syndrome. PIH is a common clinical presentation during pregnancy and requires at least one more sign such as pregnancy induced proteinuria to be considered PET. [1] Higgins and de Swiet studied different definitions used for PET by clinicians and researchers and came to a conclusion that even though many can argue whether the definition for PET is really important, it is essential to understand that clinicians and researchers seek a PET definition for different reasons. Clinicians seek to define PET to identify a group of women at higher than average risk to the health of the mothers or their babies. The researchers seek to define PET because they want to be certain that they are studying PET and not the other disease. Taking this statement into consideration, the authors concluded that any definition used clinically should be as loose as possible but any PET definition used by researchers should be stringent. [53] This study was supported by the results of the previous study by North et al. who discovered a correlation between PET definition and rate of pregnancy complications and suggested an evidence-based approach be used in defining PET. [54] The importance of research on PET/PIH cannot be overstated because of the effect of these conditions on the wellbeing of mothers and their babies not only perinatally but later in life as well. [1] But as is reported by many authors [22, 23], the research of these

conditions is hindered by a lack of consistency in the definition of PET used in obstetrical research. Consequently, the results from many studies are not comparable.

2.7 Current Definition of Hypertension in Pregnancy

According to Society of Obstetricians and Gynecologists of Canada (SOGC), [2] hypertension in pregnancy is classified as pre-existing or gestational hypertension. Pre-existing hypertension pre-dates pregnancy or appears before 20 weeks, and gestational hypertension appears at or after 20 weeks. For both pre-existing and gestational hypertension, there are two subgroups: 1) with comorbid conditions and 2) with preeclampsia, defined by three criteria: hypertension (new or worsening), proteinuria, and adverse conditions. According to SOGC Guideline, the term, preeclampsia has been re-introduced for its brevity and because of its international use. It corresponds to the following **previous** terms:

- pre-existing hypertension with superimposed gestational hypertension, proteinuria and/or an adverse condition or conditions
- gestational hypertension with proteinuria
- gestational hypertension (without proteinuria) with one or more of the adverse conditions.

The classification emphasizes that there is significant clinical overlap; women may meet criteria for more than one subgroup, and that evolution may occur over time. A final diagnosis of the type of Hypertensive Disorder of Pregnancy (HDP) is retrospective, following the postpartum period.

The most commonly used definition for PET describes it as a hypertensive disorder with new-onset proteinuria, and, potentially, other end-organ dysfunction. A restrictive definition of preeclampsia is gestational hypertension with proteinuria. The latter is often used by the research community. A more inclusive definition of PET is gestational hypertension with proteinuria or typical end-organ dysfunction.

We will use the definition of PET advanced by SOGC because, first, it takes into consideration the pathogenesis of PET as a key to understanding the multi-system and varied clinical manifestations of PET and, second, because it has been endorsed by the obstetrical, internal medicine and research community in Canada. The definition mitigates much of the confusion in previous definitions.

Here is how hypertension in pregnancy is classified in ICD-10-CA: [55]

O10.0 “Pre-existing essential hypertension complicating pregnancy, childbirth and the puerperium;”

O10.1 “Pre-existing hypertensive heart disease complicating pregnancy, childbirth and the puerperium;”

O10.2 “Pre-existing hypertensive renal disease complicating pregnancy, childbirth and the puerperium;”

O10.3 “Pre-existing hypertensive heart and renal disease complicating pregnancy, childbirth and the puerperium;”

O10.4 “Pre-existing secondary hypertension complicating pregnancy, childbirth and the puerperium;”

O10.9 “Unspecified pre-existing hypertension complicating pregnancy, childbirth and the puerperium;”

O11.0 “Pre-existing hypertension with superimposed proteinuria;”

O12.0 “Gestational (pregnancy-induced) oedema and proteinuria without hypertension;”

O13.0 “Gestational (pregnancy-induced) hypertension without significant proteinuria;”

O14.0 “Gestational (pregnancy-induced) hypertension with significant proteinuria;”

O15.0 “Eclampsia;”

O16.0 “Unspecified maternal hypertension.” (Appendix 1)

HDP grouping according to SOGC and their approximation in the ICD-10-CA is presented in Table 3.2.

Table 3.2 SOGC grouping of HDP and their approximation in the ICD-10_CA

SOGC definition	Corresponding ICD-10-CA codes
Chronic Hypertension	O10
Gestational hypertension	0.13 (Gestational, pregnancy-induced, hypertension without significant proteinuria)
PET	O11 (Pre-existing hypertensive disorder with superimposed proteinuria), O14 (Gestational, pregnancy-induced, hypertension with significant proteinuria), O15 (Eclampsia in labour)
HDP	O11 (Pre-existing hypertensive disorder with superimposed proteinuria), 0.13 (Gestational, pregnancy-induced, hypertension without significant proteinuria) O14 (Gestational, pregnancy-induced, hypertension with significant proteinuria), O15 (Eclampsia in labour)

Because the modern classification of hypertensive disorders of pregnancy (ICD-10) used in hospital discharge database is not ideal for reporting of PET, in our study we will assess the diagnostic accuracy of hospital discharge database via ICD-10-CA coding in

reporting of PET in comparison with adjudicated data from a prospective birth cohort study used as a “reference standard”.

Chapter 3 Proposed Approach

In order to fully explore the proposed topic, the project will consist of three subprojects. First, we will analyze inter-rater agreement among raters adjudicating preeclampsia cases from the Ottawa and Kingston (OaK) Birth Cohort study. The study was a prospective cohort study in which pregnant women were recruited between 12 and 20 weeks of gestation during their prenatal visit at The Ottawa Hospital and Kingston General Hospital from October 2002 to April 2009. The objective of the original cohort study was to evaluate the association between folic acid supplementation in early pregnancy and the risk of adverse placenta-mediated pregnancy outcomes including intrauterine growth restriction (IUGR), PET, placenta abruptio (PA), and fetal death.

Second, we will conduct a validation of PET diagnosis study in the hospital discharge database by comparing the frequency of PET diagnosis recorded in the discharge abstract database with the frequency of PET diagnosis in the reference standard database from the OaK Birth Cohort. We will also examine the potential causes of coding errors within the coding process and analyze predictors of misclassification.

Finally, we will examine whether and to what extent research based on hospital discharge database would distort the results by testing the association between PET risk factors and PET development in reference standard group and comparing it with the one estimated through hospital discharge database and the ICD-10-CA.

Chapter 4 Project 1

Inter-rater Agreement Estimation

4.1 Aim of the Project

To calculate the inter-rater agreement of the PET diagnosis that occurred during the outcome adjudication of the OaK Birth Cohort study data.

4.2 Hypothesis

The level of agreement with the PET diagnosis among the reviewers adjudicating the outcome data from the OaK Birth Cohort study is considered good.

4.3 Background

4.3.1 Introduction to the Concept of Bias

When designing a research, two concepts should be taking into consideration. The first concept is the internal validity of the study. It allows drawing conclusions about the broader population based on information from a study population. The other concept is the external validity or generalizability of the results of a study to the broader population. A research study should avoid bias or systematic error. Bias undermines the internal validity of research and measures deviation of the study results from results of a valid study. All observational research study (e.g. cohort, case-control) have built-in bias. [56]

There are different kinds of bias in observational studies: selection bias and information bias, as well as other biases. Information bias is a systematic error in a study that arises because of incorrect information obtained on one or more variables measured in the study. As an example, misclassification of exposure and/or disease, due to inaccurate information, can lead to a biased measure of effect. [56]

The main focus of this thesis is to evaluate the effect of diagnosis misclassification, or misdiagnosis, in particular, on the measure of effect. There are several reasons why misclassification bias might occur (depending on the source from which information was obtained). The main reasons are:

- a) Incorrect diagnosis that results from:
 - Limited knowledge about the disease
 - The complex diagnostic process
 - Subclinical disease
 - Detection bias (e.g., a physician gives a more thorough exam to a patient who has more severe condition)
- b) Subject self-report records incorrectly coded in the database [54]

4.3.2 Quantitative Assessment of the Information Bias

The information bias can be evaluated by building a misclassification table that provides a convenient summary of how disease status can be misclassified from a true disease status to an observed disease status. [56] From this table different parameters can be calculated such as:

- Sensitivity, the probability that a subject who is truly diseased will be classified as diseased in one's study,
- Specificity, the probability that a subject who is truly non-diseased is classified as non-diseased in one's study,
- Positive Predictive Value (PPV), the probability of actually having the disease when the test is positive,

- Negative Predictive Value (NPV), the probability of actually not having the disease when the test is negative.

4.3.3 What Can Be Done to Decrease Information Bias?

Adjudication is a recommended process to standardize the outcome assessment for randomized clinical trials (RCT) even though RCTs, as a study design, are less susceptible to bias than any other study designs. [57, 58] Some studies have shown the end-point adjudication can correct the number of end-point events and, therefore, a significant difference in the study results interpretation has been achieved. [59, 60] However, the other studies, that have evaluated the effect of the outcome adjudication on the result estimates of the outcome, failed to demonstrate any significant effect of event adjudication on the study conclusions. [61] Overall, the results of the studies evaluating the effect of the outcome adjudication on the estimates for treatment effects in RCTs are inconclusive.

RCTs are more rigorous and less susceptible to different kinds of bias study design than observational studies; thus one would expect that adjudication of the event would be used more often in observational research to decrease probability of information bias. However, it is not the case in majority of observational studies. Furthermore, the presentation of the results of adjudication differs from one study to another; some studies present results in percentage of disagreement but other studies present agreement as inter-rater agreement kappa. [62] We found only a few studies that used adjudication process in observational research and none in obstetrics. In this first part of the thesis, we evaluated the inter-rater reliability of three experts adjudicating the diagnosis of PET.

4.3.4 Kappa Coefficient (Cohen's Kappa)

4.3.4.1 Definition of the Agreement

The agreement between the judgments is defined as the conformity of two or more information relating to the same object. The statistical analysis of observer agreement in diagnosis is generally performed for three reasons. First, observer agreement provides information about the reliability of a diagnosis. A reliable method should produce a good agreement when used by knowledgeable observers. Second, observer agreement can be used to check the consistency of a method for classification of an abnormality that indicates the extent or severity of disease, and to determine the reliability of various signs of the disease. It can also be used to compare the performance of humans and computers. Third, observer agreement can provide a general estimate of the value of a diagnosis technique when an independent method of proving the diagnosis precludes the measurement of sensitivity and specificity or the more general receiver operating characteristic curve. [63]

The Kappa coefficient K was proposed by Cohen [63] in 1960 to quantify the intensity or quality of the real agreement between paired qualitative judgments. K of a maximum percentage of agreement corrects what would be under the simple effect of chance. The true value of the Kappa coefficient in the population is a random variable approximately follows a Gaussian law with mean K and variance $\text{Var}(K)$. The null hypothesis H_0 is $K = 0$ against the alternative hypothesis $H_1: K > 0$. In the case of a study of the agreement between two statistically independent observers with r terms (categories) trial, with $r \geq 2$, the Kappa coefficient, K , is written:

$$K = \frac{P_o - P_e}{1 - P_e} \quad (4.1)$$

Where P_o is the proportion of agreement observed and P_e is the proportion of random agreement or concordance expected under the assumption of independent judgments. Table 4.1 presents the notation used when data are presented in a contingency table.

Table 4.1 Proportions Joined Judgments of Two Judges on a Scale with r Categories

		judge A				
Category	1	2	...		r	Total
judge B	1	$p_{11} \ p_{12} \ \dots$			p_{1r}	$p_{1.}$
	2	$p_{21} \ p_{22} \ \dots$			p_{2r}	$p_{2.}$
	r	$p_{r1} \ p_{r2} \ \dots$			p_{rr}	$p_{r.}$
Total		$p_{.1} \ p_{.2} \ \dots$			$p_{.r}$	1

To compute Kappa for 2 raters, first we need to calculate the observed level of agreement

$$P_o = P_{11} + P_{22} \quad (4.2)$$

This value needs to be compared to the value that you would expect if the two raters were totally independent,

$$P_e = P_{1.}P_{.1} + P_{2.}P_{.2} \quad (4.3)$$

The value of Kappa, K , is defined as shown in equation (4.1).

The numerator represents the discrepancy between the observed probability of success and the probability of success under the assumption of an extremely bad case. Independence implies that pair of raters agrees about as often as a pair of people who effectively flip coins to make their ratings. The Kappa coefficient is a real dimensional number between -1 and 1. The agreement will be higher when the value of Kappa is close to 1 and the maximum agreement is reached ($K = 1$) where $P_o = 1$ and $P_e = 0.5$. When there is independence of judgments, the Kappa coefficient is equal to zero ($P_o = P_e$), and in the case of a total disagreement between the judges, the Kappa coefficient takes the value -1 with $P_o = 0$ and $P_e = 0.5$. It is possible for Kappa to be negative, but this does not occur too often. In such a case, one should interpret the value of Kappa to imply that there is no effective agreement between the two rates. [64] Only the positive values of Kappa measure the level of agreement among raters while zero or negative values indicate independence of judgements or disagreement respectively. When multiple readers are used, it was suggested by some authors to calculate the values of Kappa for pairs of readers and then compute an average Kappa for all possible pairs. [65, 66]

Landis and Koch have proposed a classification of the agreement based on the value of Kappa [66] (Table 4.2.)

Table 4.2 Level of Agreement and Kappa Value

<i>Agreement</i>	<i>Kappa</i>
Excellent	≥ 0.81
Good	0.80 to 0.61
Moderate	0.60 to 0.41
Poor	0.40 to 0.21
Bad	0.20 to 0.0
Very bad	<0.0

4.3.4.2 Agreement between Several Judges

The identity between the correlation coefficient and intraclass Kappa coefficient is operated by Fleiss [65] to formulate the kappa statistics using a variance analysis of information obtained by coding the verdicts and judgments in two categories, positive and negative.

Let m judges (≥ 2), who independently are asked to rate n sample observations into two categories of judgments: positive and negative. It defines the proportion of rank n observations by m judges in the positive category, including:

$$\bar{P} = \frac{\sum_{i=1}^n x_i}{n\bar{m}} \quad (4.4)$$

Where x_i is the number of judgments in the positive category for observation i and

$$\bar{m} = \frac{\sum_{i=1}^n m_i}{n} \quad (4.5)$$

which is the average number of judgments by observation and m_i the total number of judgments for observation i .

The average of squared deviations between observations (BMS) is approximately equal to:

$$BMS = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - m_i \bar{P})^2}{m_i} \quad (4.6)$$

if $n \geq 20$.

The average of squared deviations intra-observer (WMS) is equal to:

$$WMS = \frac{1}{n(\bar{m}-1)} \sum_{i=1}^n \frac{x_i(m_i - x_i)}{m_i} \quad (4.7)$$

The coefficient of intraclass correlation is estimated by:

$$r = \frac{BMS - WMS}{BMS + (m_0 - 1)WMS} \quad (4.8)$$

with

$$m_0 = \bar{m} - \frac{\sum_{i=1}^n (m_i - \bar{m})^2}{n(n-1)\bar{m}} \quad (4.9)$$

If the number of observations n is large, m_0 and the average number of judgments per case ($\bar{m}$) are somewhat different since they are equal if the number of judgments is constant for all observations. If we replace m_0 by $\bar{m}$, the expression of the correlation coefficient and intraclass Kappa become:

$$K = \frac{BMS - WMS}{BMS + (\bar{m} - 1)WMS} \quad (4.10)$$

this is equivalent to write:

$$K = 1 - \frac{\sum_{i=1}^n \frac{x_i(m_i - x_i)}{m_i}}{n(\bar{m} - 1)\bar{p}\bar{q}} \quad (4.11)$$

with $\bar{q} = 1 - \bar{p}$.

Intraclass correlation evaluates the level of agreement between raters in measurements, where the measurements are parametric or at least interval. This method is better than ordinary correlation as more than 2 raters can be included, and there is a correction for correlations between raters that becomes apparent when the range of measurement is large. The coefficient represents concordance, where 1 is perfect agreement and 0 is no agreement at all. [64]

4.4 Methods

4.4.1 Population

OaK Birth Cohort study

8085 mothers had been recruited in the OaK Birth Cohort recruited from October 2002 to April 2008 and followed until birth. The objective of the original cohort study was to evaluate the association between folic acid supplementation in early pregnancy with the risk of adverse placenta-mediated pregnancy outcomes including intrauterine growth restriction (IUGR), PET, placenta abruptio (PA), and fetal death.

To be included in the study, women had to be between 12 and 20 weeks of gestation, with a viable fetus, and an expected singleton or twin pregnancy. Women were recruited by experienced obstetrical research nurses who explained the purpose of the study and

what participation involved. The following data was obtained for each participant of the OaK study:

- Demographic data (date of birth, weight, height, income, education, ethnicity);
- Past pregnancy history (history of PET, PA, fetal death, or gestational diabetes in previous pregnancies);
- Pre-existing medical conditions (history of chronic hypertension, diabetes, or other chronic conditions);
- Social History (smoking, alcohol consumption during pregnancy);
- Delivery information (date of birth of the baby, gestational age at birth, mode of delivery, status of the baby at birth, weight of the baby and other variables);
- Pregnancy complications (PA, HDP/PET).

Demographic and health information was collected using a case report form (CRF) administered by the obstetrical research nurses. Although demographic, past pregnancy history, social history and pre-existing medical conditions data were self-reported by the study participants, this information was also verified against information from the participant medical records to ensure the accuracy of the data. Pregnancy outcomes data were later collected from hospital charts after delivery to complete the CRF for each participant. Infant data, such as sex, weight, and other variables were also collected at birth. A total of 5277 women were recruited from the Ottawa Hospital and 2808 from the Kingston General Hospital. Recruitment statistics at the Ottawa Hospital as following:

1070 women were recruited from the Civic campus, 86 women from the Riverside campus, and 4121 women were recruited from the General campus of the Ottawa hospital. Due to logistical reasons the majority of the OaK study participants were recruited at the General campus of the Ottawa Hospital. The majority of the participants delivered at the Ottawa Hospital General campus, followed by the Ottawa Hospital Civic campus. The minority of the participants delivered in other hospitals of city of Ottawa (Montfort, Queensway Carleton) and elsewhere (other provinces). The exact distribution of participants based on the hospital of delivery was unknown. Only participants who delivered in the Ottawa Hospital (General and Civic Campuses) were included into our study because only administrative data from this hospital required for comparison were accessible.

4.4.2 Outcome Definition

The following definition was used for the study outcome, PET (according to SOGC):

PET was defined as new-onset hypertension and proteinuria or one or more of the other adverse conditions.

Hypertension in pregnancy was defined as a diastolic BP of ≥ 90 mmHg, based on the average of at least two measurements.

Proteinuria was strongly suspected when urinary dipstick proteinuria is $\geq 2+$. Proteinuria was defined as ≥ 0.3 g/d in a 24-hour urine collection or ≥ 30 mg/mmol urinary creatinine in a spot (random) urine sample.

The patient was classified, due to adverse conditions, as having PET even if she did not

meet the criteria for proteinuria. [2] Among these adverse conditions were maternal symptoms, maternal signs of end-organ dysfunction and abnormal maternal laboratory testing). Maternal symptoms included persistent or new/unusual headache, visual disturbances, persistent abdominal or right upper quadrant pain, severe nausea or vomiting, and chest pain or dyspnea. Maternal signs of end-organ dysfunction included eclampsia, severe hypertension, pulmonary edema, or suspected placenta abruptio. Abnormal maternal laboratory testing included elevated serum creatinine, aspartate aminotransferase (AST), alanine aminotransferase (ALT), lactate dehydrogenase (LDH) and/or platelet count less than $100 \times 10^9/L$ (according to local laboratory criteria) along with hypertension.

Severe preeclampsia corresponded to preeclampsia with onset before 34 weeks of gestation, with heavy proteinuria (3–5 g/d according to other international guidelines [13]), or with one or more adverse conditions. [2]

Chronic or pre-existing hypertension was defined as documented hypertension that developed before pregnancy or appeared before 20 weeks. [2] Chronic hypertension was diagnosed if the patient had documentation of chronic/essential hypertension in the chart.

Gestational hypertension (GH) was diagnosed if the patient met SOGC criteria for hypertension but did not have significant proteinuria or other signs/symptoms for PET.

Hypertension during Pregnancy (HDP) was defined as hypertension developed at or after 20 weeks which included GH and PET. [2]

All available charts were reviewed and a short synopsis for every case was prepared. All cases, with HDP/PET and without it, were presented to the adjudication committee for reviewing. The results of adjudication included:

1. Normal cases (women who never fulfilled criteria for HDP/PET)
2. Chronic hypertension cases;
3. PET cases;
4. HDP cases;

4.4.4 Data collection

Data were collected by using study CRFs. Extra form was created to incorporate additional information from the study participant's medical chart to the case report forms (see Appendix 2).

4.4.5 Process for event adjudication

Figure 4.1 summarizes the adjudication process that we are going to discuss here in more details. The adjudication committee included an obstetric internist, a perinatal epidemiologist and a maternal-fetal medicine physician. The three adjudicators were involved in extensive PET research. The composition of the panel remained unchanged throughout the study. The process was designed for broad identification of all participants with possible PET.

Each case of suspected PET was identified from a CRF used to collect data for the OaK study. Additional information obtained from supporting documents found in the medical charts (delivery information, antenatal record, physician notes and lab work) was collected in the extra form for PET adjudication. This form summarizes clinical events, lab work, and other pregnancy outcome information. Each case was presented to the members of adjudication group and was reviewed independently in accordance with the

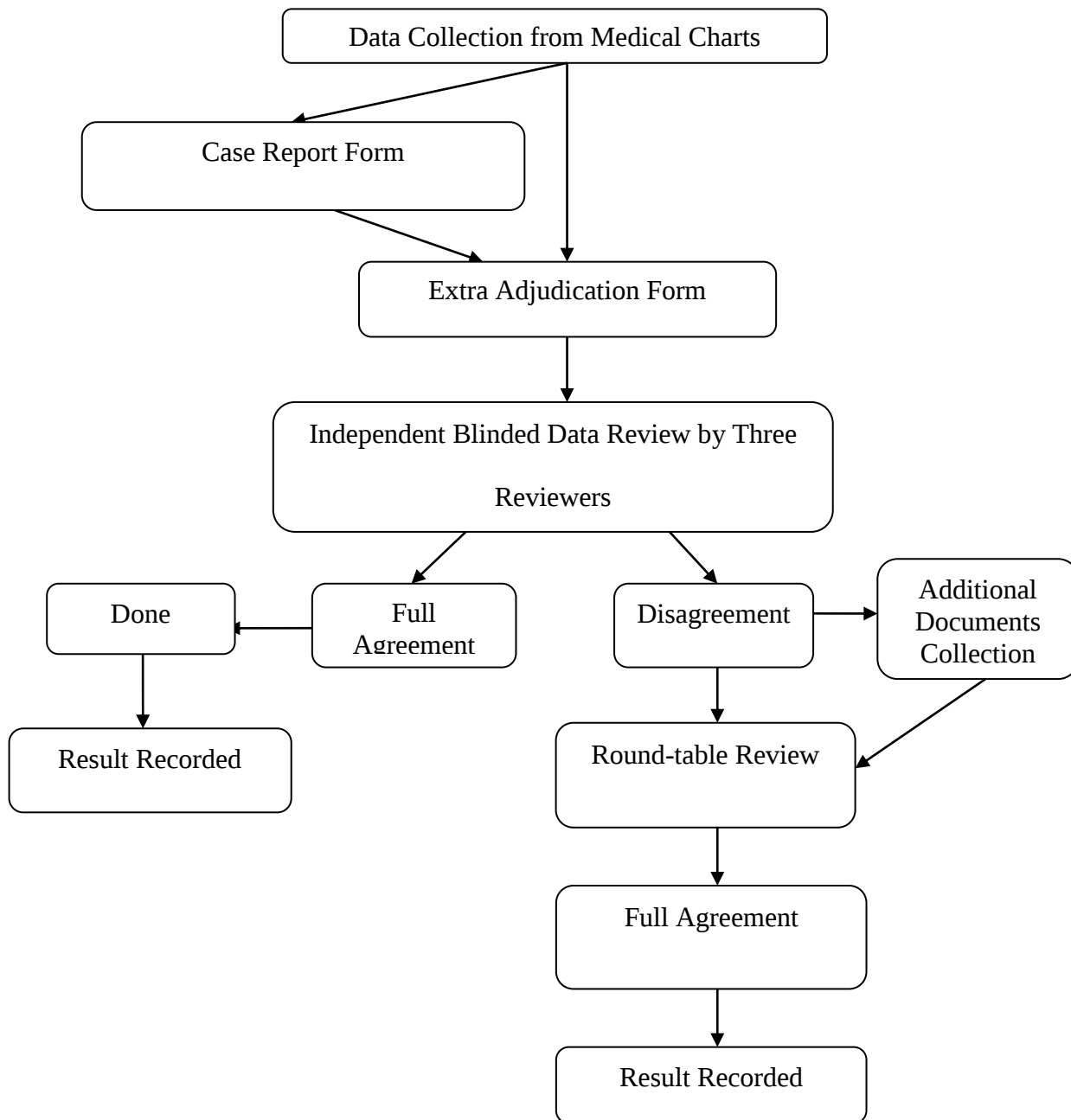
current clinical practice guideline. [2] The results of the individual adjudications were collected by a study coordinator and divided into two groups: those in which the unanimous agreement has been reached (all three reviewers had adjudicated the case as a PET case) and those in which there was disagreement (for example, two out of three reviewers adjudicated the case as a PET case, but one reviewer had judged it as a GH case). The cases with full agreement of the reviewers were entered into the outcome database. The cases with disagreement were further discussed at a round-table meeting of the reviewers until full consensus on the diagnosis was achieved, and then entered into the database. After the last case was adjudicated, the inter-rater agreement kappa was calculated.

The underlying causes for disagreement were explored. The hypotheses for the underlying reasons for disagreement were the conservative versus liberal approach to the diagnosis of PET by the reviewers and severity of PET (i.e., we hypothesised that it was easier to diagnose severe cases of PET than the mild ones).

We further calculated the values of Kappa for each pair of readers and then computed the average Kappa for all possible pairs.

A kappa that is greater than 0.75 was set to represent good agreement.

Figure 4.1 Adjudication Process



4.5 Adjudication Results

From 2887 participants of OaK birth cohort study 205 cases with suspected PET were identified and adjudicated. The inter-rater agreement was calculated.

Kappa was equal 0.82 (see Appendix 3 for the calculations).

Overall, the level of agreement on diagnosis of PET among the raters was excellent beyond that expected by chance alone.

4.6 Discussion

To our knowledge, this is the first study that has used adjudication of an end-point event in observational research in obstetrics. Even though observational research is susceptible to different kinds of bias as discussed earlier, information bias being one of them, we did not find any observational study in obstetrical research that used adjudication to reduce this kind of bias. There was an overall agreement of 0.82 which represents an excellent level of agreement. Our results show that the disagreement among reviewers on the PET diagnosis was present in 18% of cases, leaving nearly one out of five cases for additional review and discussion.

These results are not surprising. PET diagnosis belongs to the high risk category, following the categorization of the reasons for disease misclassification, because it belongs to subjective primary outcomes, including non-fatal events. [66] As discussed previously, PET is a disease with unclear etiology and pathogenesis. Different classifications have been used by different clinicians and researchers. The classification usually depends on the severity of the condition (severe cases are easily diagnosed while less severe are more prone to being misdiagnosed). [67] We observed a similar tendency

during the process of adjudication. The cases with disagreement tended to be milder ones without enough documents to support the diagnosis. We did not find that the result of adjudication depended on the approach the reviewer used (conservative vs liberal).

4.7 Conclusion

We conclude that the adjudication of a subjective primary outcome offers a way to reduce information bias in outcome ascertainment in observational research.

Chapter 5 Project 2

Validity of PET diagnosis in an administrative database

5.1 Aim

1. To analyze sensitivity, specificity, PPV, NPV of ICD-10-CA coding system in reporting PET diagnoses in the administrative database (hospital discharge database) in comparison with the “reference standard” which is the adjudicated data from a prospective birth cohort study;
2. To explore potential causes of coding errors within coding process and predictors of misclassification.

5.2 Hypotheses

1. A sizable proportion of the PET diagnosis retrieved from the hospital discharge database by using ICD-10-CA code could be misclassified.
2. There could be differences in the characteristics between misclassified cases versus correctly classified cases.
3. Chart review may be able to provide explanations for the PET misclassification.

5.3 Background

5.3.1 Ways to Measure Code Accuracy

There are five statistics commonly used to measure the accuracy of coding: sensitivity, specificity, PPV, NPV, and kappa coefficient. Sensitivity and specificity are often used when a particular “gold standard” is available. Kappa statistics are used when diagnostic labels assigned by two or more experts (e.g., physicians) evaluating the same sample of

patients. This statistics quantifies beyond-chance agreement among experts. In case the researchers want to evaluate how well medical coders' ICD-code assignments match those of physicians, the physicians' review is considered as an appropriate "gold standard" and sensitivity, specificity, and predictive values are calculated. [68]

5.3.2 Reference Standard Requirements

In performing diagnostic research of the outcome variable, the presence or absence of the target disorder must be measured and compared with the gold standard to estimate test accuracy. A real gold standard test should have 100% sensitivity and 100% specificity which is rare in real world, therefore, the term "reference standard" is more appropriate. The test to establish the presence of the disease should fulfil the following requirements to be considered a reference standard:

- It should be consistently applied for all included subjects;
- The reference standard result should be established without knowing the result of the test, in other words, blinding should be maintained to avoid "diagnosis review bias";
- The reference standard should be performed and interpreted by using standard criteria which are especially important when the standard diagnosis depends on subjective interpretations because in such case inter-observer variability can occur. [69] In studies assessing accuracy of administrative database, the physician diagnosis, as recorded in the medical charts, was considered the gold standard for measuring accuracy. However some researchers doubted that the physician diagnoses recorded in the medical charts can be considered a gold standard. [70]

5.3.5 Research Assessing Validity of PET Diagnosis in Administrative Databases

Even though administrative databases have been widely used for different purposes, there is a very limited experience regarding the validity of PET diagnosis in such databases.

There are two types of studies in the literature focusing on evaluating the quality of administrative data:

1. Auditing studies when experienced coders independently reassign diagnosis codes to a random sample of hospital records while adhering to coding standards;
2. Validation studies which compare administrative data with other sources of clinical data such as data from chart review. [71]

Researchers reported that the accuracy of data derived from administrative database for PET research is an issue. [67, 72, 73] Thornton et al. compared reporting of hypertensive disorders of pregnancy from an administrative database, coded by ICD-10-AM, with disorder-specific database (the Hypertensive Disorders of Pregnancy Database was maintained prospectively for the months of July to December 2002) which was considered a gold standard. They discovered that ICD coding system for hypertensive disorders resulted in an overall error rate of 64% and they came to a conclusion that current administrative database systems are not reliable for recording maternal medical conditions such as hypertension, and the accuracy of the reporting for this particular disorder can be achieved only with the use of disorder-specific database. The misrepresentation of maternal medical conditions may have a significant implication for clinical reporting and research activities. [73] Geller et al. conducted a study to determine the accuracy of ICD-9 codes for PET and eclampsia as defined by ACOG 2002. The authors showed that ICD-9 codes for PET varied greatly in their accuracy of diagnosis.

They found that overall PPV for severe PET, mild PET and eclampsia combined was only 54%; however PPV for severe PET only was considerably higher than for mild PET and eclampsia. [75] Authors reported that the most common error was diagnostic error followed by coding error, and these two kinds of errors accounted for 82.1% of all cases. Authors concluded that current ICD-9 code alone cannot be used to monitor trends of PET, and suggested that additional training for physicians and coders may improve the accuracy of the database. [74] Klemmensen et al. conducted a validation study of accuracy of administrative databases and ICD-10 codes for reporting of hypertensive disorders in pregnancy by comparing Danish national hospital discharge registry with the chart review from three other hospitals which was considered “the gold standard”. Authors used ACOG 2002 definition for PET in this study. They reported low sensitivities for all types of hypertensive disorders (69.32 %) and for individual subtypes but high specificities for all hypertensive disorders (99.47%). The authors also reported that the incidence of PET derived from the administrative database was similar to the one from the chart review (2.72 and 2.90% respectively); however gestational hypertension was severely underestimated in the administrative database. Based on the results the authors concluded that research based on administrative data via ICD-10 coding was satisfactory for etiologic studies where preeclampsia was a primary outcome. [76] But other researchers such as Callaghan doubted that high specificity of administrative data and ICD-10 coding of hypertensive disorders of pregnancy was sufficient for a group of disorders with unknown pathophysiology and lack of specific biomarkers even if used only for etiologic research. [77] Roberts et al. performed a validation study assessing the accuracy of the reported hypertensive disorders in pregnancy in two administrative

databases (birth database and hospital discharge database) by comparing them with the data abstracted from medical records which was considered the “gold standard”. [74] Their findings were similar to Klemmensen et al. [75] in that pregnancy induced hypertension and chronic hypertension reported by administrative database had a high specificity but low sensitivity. Using more than one source of data may improve accuracy. They also discovered that mild cases of PET, without significant adverse outcomes, had higher probability of being underreported than severe cases. Another conclusion from this study was that, because ICD (ICD-10) did not provide definition of the hypertensive disorders of pregnancy, it led to significant misclassification of gestational hypertension and PET, which could be improved by collapsing these categories into a broader category such as pregnancy hypertension; however this could have limited clinical utility. [74] Joseph and Mahey studied validity of data from the hospital discharge in comparison with data from the Nova Scotia Atlee Perinatal Database which was considered a gold standard. Among other findings, sensitivity for gestational hypertension with proteinuria was 75.2% and 87.9% for any gestational hypertensive disorder with specificity 99.5 and 99.6%, respectively. [79]

Klemmensen et al. suggested that because of the lack of consistency in guidelines regarding PET, medical personnel who report hypertensive disorders in pregnancy may encounter difficulties coding this condition. They concluded that it is essential to have comprehensive national guidelines agreed upon and followed by obstetricians. They also stated that national guidelines need to be similar to the international ones to enable comparison of studies. [78]

In conclusion, most studies that evaluated the accuracy of administrative databases via

the ICD coding system for reporting hypertensive disorders in pregnancy described significant misclassification of the conditions.

5.3.6 The sources of coding errors

According to O'Malley et al. main sources of coding errors in relevance to the patient included:

- amount and quality of information at admission,
- communication between patients and providers,
- the clinician's knowledge and experience with the illness, and
- the clinician's attention to detail.

Main sources of coding errors in relevance to the medical records among others included:

- variance in coder training and experience,
- facility quality control guidelines, and
- variance in amount of information available and definitions/terminology used.[80]

Cheng et al. conducted a chart review audit to evaluate discrepancies in clinical codes assignment and underlying causes of coding errors. They concluded that while high-level skills in addition to depth and currency of knowledge in clinical coding are critical for good coding, the dominant need for improvement lies in the documentation. [31] As a direct outcome of the study, the problems with the documentation have been addressed via the following actions, based on the auditor's recommendations:

- A. The internal clinical coding auditor meets with senior clinicians from each clinical department to address documentation problems and answer queries, for example

relating to suspected conditions or principal diagnosis selection, which may impact the coding practice.

- B. The internal auditor attends clinical case study meetings to advise on the relationships between coding requirements and medical record documentation.
- C. The Coding Unit Manager and/or Coding Educator conduct education programs for interns and other junior clinicians on medical record documentation.
- D. Health information managers from the hospital's Health Information Service provide education to clinicians on case-mix and DRGs, to help them understand the importance, in a case-mix-based funding environment, of accuracy and completion of their medical record documentation.

Sources of coding errors were identified and it was concluded that lack of communication between coders and clinicians is one of these sources due to following factors:

- clinicians do not feel the fact that the coded data reflect their work
- coders feel intimidated about asking questions, seeking advice or asking about clinical issues

Different ways to improve the communication between clinicians and coders were recognized.

5.4 Methods

5.4.1 Data sources

5.4.1.1 OaK Birth Cohort

OaK Birth Cohort study

8085 mothers had been recruited in the OaK Birth Cohort recruited from October 2002 to

April 2008 and followed until birth. The objective of the original cohort study was to evaluate the association between folic acid supplementation in early pregnancy with the risk of adverse placenta-mediated pregnancy outcomes including intrauterine growth restriction (IUGR), PET, placenta abruptio (PA), and fetal death.

To be included in the study, women had to be between 12 and 20 weeks of gestation, with a viable fetus, and an expected singleton or twin pregnancy. Women were recruited by experienced obstetrical research nurses who explained the purpose of the study and what participation involved. The following data was obtained for each participant of the OaK study:

- Demographic data (date of birth, weight, height, income, education, ethnicity);
- Past pregnancy history (history of PET, PA, fetal death, or gestational diabetes in previous pregnancies);
- Pre-existing medical conditions (history of chronic hypertension, diabetes, or other chronic conditions);
- Social History (smoking, alcohol consumption during pregnancy);
- Delivery information (baby's date of birth, gestational age at birth, mode of delivery, baby's status at birth, baby's weight and other variables);
- Pregnancy complications (PA, hypertension in pregnancy).

Demographic and health information was collected by using a CRF administered by the obstetrical research nurses. Although demographic, past pregnancy history, social history, pre-existing medical conditions data were self-reported by the study participants, the information was also verified against information from the participant medical records to

ensure the accuracy of the data. Pregnancy outcomes data were later collected from hospital charts after delivery to complete the CRF for each participant. Infant data, such as sex, weight, and other variables were also collected at birth. A total of 5277 women were recruited from the Ottawa Hospital and 2808 from the Kingston General Hospital. Recruitment statistics at the Ottawa Hospital as following:

1070 women were recruited from the Civic campus, 86 women from the Riverside campus, and 4121 women were recruited from the General campus of the Ottawa hospital. Due to logistical reasons a majority of the OaK study participants were recruited at the General campus of the Ottawa Hospital. A majority of the participants delivered at the Ottawa Hospital General campus, followed by the Ottawa Hospital Civic campus. A minority of the participants delivered in other hospitals of city of Ottawa (Montfort, Queensway Carleton) and elsewhere (other provinces). The exact distribution of participants based on the hospital of delivery was unknown. Only participants who delivered in the Ottawa Hospital (General and Civic Campuses) were included into our study because only administrative data from this hospital required for comparison were accessible.

Adjudicated data from the OaK Study was considered as a “reference standard” for this project because every case of HDP/PET was adjudicated by a group of experts blinded to the ICD-10-CA code and in accordance with the current SOGC Guidelines for HDP. [2]

5.4.1.2 Administrative database

Discharge summary information was obtained from the Ottawa Hospital discharge database for a randomly selected 2890 participants of the OaK Birth Cohort study.

5.4.2 Sample size calculations

Calculation of the sample size was based on sensitivity of the hospital discharge database to diagnose PET obtained from the literature review. [81] According to the review, the sensitivity was 70%, the prevalence of PET was 5%, [1, 2] and the minimal acceptable level of confidence interval was 0.55 (which was set at 0.95) (Flahault et al. 2004). [82]

The minimal sample size of cases was set as 114 as presented in Table 7 in Appendix 4.

The corresponding number of controls was obtained from the following equation:

$$N_{\text{controls}} = N_{\text{cases}} [(1 - \text{Prev})/\text{Prev}] = 114 [(1 - 0.05)/0.05] = 2736$$

As a result, it was necessary to have 114 PET cases and 2736 non-PET cases (controls) (see Appendix 4 for details).

We re-calculated the sample size based on the actual number of cases (103) and incidence of PET from our study population to obtain the number of controls required to achieve the appropriate sample size to power the study.

$$N_{\text{controls}} = N_{\text{cases}} [(1 - \text{Prev})/\text{Prev}] = 103 [(1 - 0.036)/0.036] = 2758$$

We added this number to the number of cases: $103 + 2758 = 2861$. We concluded that the sample from the original cohort was sufficient to satisfy the power need.

5.4.3 Design

Retrospective cohort study

5.4.4 Data sampling

Random sample of 2890 charts from participants of OaK study delivered in Ottawa Hospital was created using SAS, Windows version 9.2 (SAS Institute, Cary, NC, USA).

5.4.5 Data linkage

Participants of OaK Study delivered in the Ottawa Hospital were linked by their hospital chart numbers and discharge date (to avoid the confusion in case there were multiple entries for the same patient) with the Ottawa Hospital discharge database in order to obtain discharge summaries and ICD-10-CA codes [53] and all charts for the sample study subjects were requested from Medical Records Department (separately for Civic and General campuses). Following cases with corresponding ICD-10-CA codes were identified in the Ottawa hospital discharge database:

1. Chronic hypertension cases included those coded as O10 code in ICD-10-CA;
2. HDP cases included codes in ICD-10-CA:
 - O11 (Pre-existing hypertensive disorder with superimposed proteinuria),
 - O13 (Gestational, pregnancy-induced, hypertension without significant proteinuria),
 - O14 (Gestational, pregnancy-induced, hypertension with significant proteinuria), and
 - O15 (Eclampsia in labour) in ICD-10-CA;
3. PET cases included the following codes in ICD-10-CA
 - O11 (Pre-existing hypertensive disorder with superimposed proteinuria),
 - O14 (Gestational, pregnancy-induced, hypertension with significant proteinuria), and
 - O15 (Eclampsia in labour) in ICD-10-CA (see Appendix 2 for ICD-10-CA codes).
4. Normal cases (women who neither fulfilled HDP/PET criteria nor coded as such)

5.4.6 Definitions

Every case with an outcome from “reference standard” database was compared with the results obtained through the Ottawa hospital discharge database. Even though the objective of the project was to assess the accuracy of hospital discharge database (via ICD-10-CA) for reporting of PET, we additionally assessed the accuracy of hospital discharge database for reporting of HDP in order to compare reliability of the hospital discharge database for reporting of both conditions. Patients were classified as true cases if they were classified as PET/HDP in the “reference standard” database and in the Ottawa hospital discharge database. The main focus was an overall summary estimate of accuracy of HDP/PET reporting expressed by its sensitivity and specificity, PPV, and NPV with their 95% confidence intervals (see Table 5.1). [83, 84]

Table 5.1 Two-by-two Table Used for the Ottawa Hospital Discharge Database Accuracy Calculations

HDP/PET	Reference standard	
Ottawa hospital discharge database	Adjudicated as “yes”	Adjudicated as “no”
Coded as “yes”	(a) True positive	(b) False positive
Coded as “no”	(c) False negative	(d) True negative

Sensitivity denotes how completely the Ottawa hospital discharge database identified PET or HDP state compared with the “reference standard”.

The specificity denotes correct ascertainment of non-hypertensive (in the case of ascertainment of HDP) or non-preeclamptic (ascertainment of PET only) state in the Ottawa hospital discharge database compared with the “reference standard”.

The PPV denotes how accurately the Ottawa hospital discharge database identifies HDP or PET and NPV for the absence of HDP or PET compared with the “reference standard”.

If a coding error occurred, the source of that error was recorded.

5.4.7 Sources of errors (PET diagnosis only)

We analyzed the source of errors based on whether diagnostic error occurred because:

- a coder recorded an incorrect ICD-10-CA code, or
- a treating physician documented an incorrect diagnosis.

We adopted the following classification of errors: [72]

- PET pregnancies coded with incorrect ICD-10-CA code;
- Normal pregnancy outcome coded as having PET;
- Pregnancies affected by PET coded as normal.

Errors rates were calculated and compared with the “reference standard”.

Potential causes of coding errors within coding process were examined.

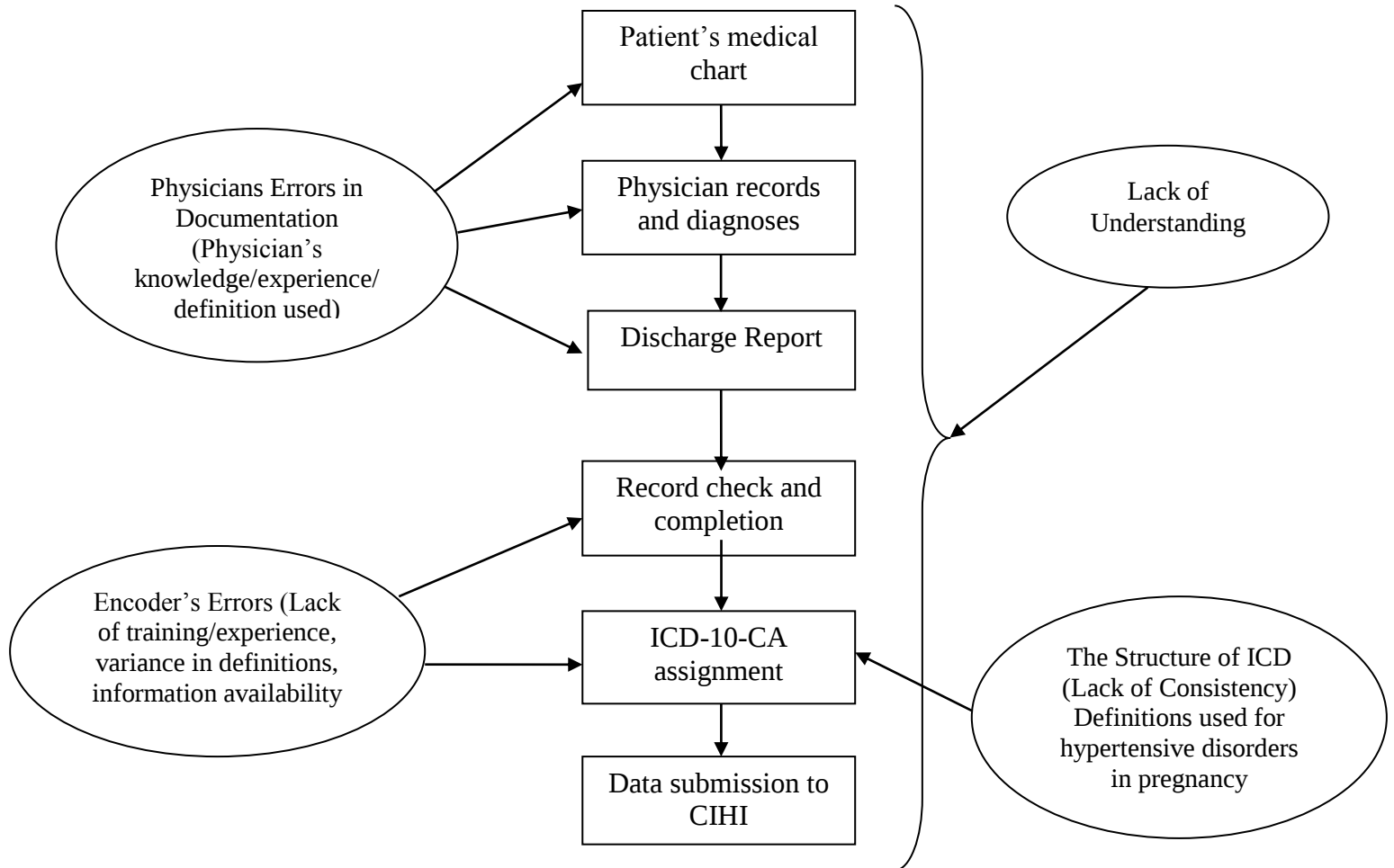
Figure 5.1 shows potential causes of coding errors within the framework of the coding process. [85]

5.4.8 Data Review

A detailed chart review of the misclassified PET cases (positive cases according to the “reference standard” but negative cases according to the Ottawa hospital discharge database) or reversed was performed to determine the presence/absence of key PET

criteria including BP readings, dipstick protein reading, 24 hour protein reading, and clinical symptoms specific to PET. All available data, beside all diagnoses recorded in the medical charts related to PET, were entered into database created specifically for this purpose.

Figure 5.1 Causes of Coding Errors



5.4.9 Variables of Interest

In order to assess characterization of misclassified cases, further analysis of different variables of interest was performed. We hypothesized that there were certain demographic and socioeconomic factors that could influence the outcome misclassification. The following variables were included into analysis:

- Maternal age,
- Maternal pre-pregnancy BMI,
- Education,
- Family income,
- Ethnicity,
- History of complications in previous pregnancies (i.e., PET, PA, or fetal death (stillbirth)),
- History of Caesarean Delivery (CS),
- Parity,
- Hospital of delivery (General vs. Civic),
- Year of delivery (2002-2006 vs. 2006-2009),
- Severity of PET according to current SOGC Guidelines for pre-eclampsia*.

*Severe preeclampsia corresponds to preeclampsia: with onset before 34 weeks' gestation, with heavy proteinuria (3–5 g/d), or with one or more adverse conditions. [2]

5.4.10 Statistical Analysis

Descriptive statistics were calculated for all socio-demographic and clinical characteristics using means and standard deviations for continuous variables and frequency counts and percentages for categorical factors. Comparisons between groups were performed using cross-tabulations and either Fishers' exact or chi-square tests of

significance. All analyses for logistic regression modelling were performed using SAS (version 9.2 SAS Institute Inc, Cary, NC).

5.4.11 Ethical Considerations

The Ottawa Hospital Research Ethics Board approval was sought for this project and obtained in January 2010. This research was carried out in accordance with the Tri-Council Policy Statement – ethical conduct for research involving humans. [86] Data were presented in aggregate form and individual patients were not identified and never contacted. (Appendix 5)

5.5 Results

5.5.1 Overall frequencies

Of the 2980 records 2887 were available for review (96.88%, others were not available at the time due to different reasons), 519 from Civic campus and 2368 from General campus of Ottawa Hospital. There is no reason to assume that the missing charts were intentionally selected, and we do not expect that they were to have any substantial impact on the results of the study. Only live and stillbirths' cases ≥ 20 weeks of gestation were included in the final calculation (terminations either spontaneous or therapeutic were excluded from the analysis based on the fact that according to the definition, PET develops after 20 weeks of gestation). 2866 records were included in the final analysis.

The disease frequencies according to the reference standard and the hospital discharge database are presented in Table 5.2 and Table 5.3 respectively.

Table 5.2 Hypertensive Diseases of Pregnancy Frequencies according to the Reference Standard Data

	Reference standard (#)	Reference standard (%)
PET (including HELLP and eclampsia)	103	3.6
GH	147	5.13
Chronic hypertension	42	1.47
Normal	2574	89.77

Table 5.3 Hypertensive Diseases of Pregnancy Frequencies according to the Ottawa Hospital Discharge Database

	Ottawa Hospital discharge database (#)	Ottawa Hospital discharge database (%)
PET (including HELLP and eclampsia)	42	1.47
GH	157	5.48
Chronic hypertension	19	0.66
Normal	2648	92.39

5.5.2 Sensitivity, Specificity, Positive Predictive Value and Negative Predictive Value

Below, we presented the results comparing PET results from the “reference standard” with presence or absence of PET codes from the Ottawa hospital discharge database coded via ICD-10-CA in a traditional two-by-two table (Table 5.4).

Table 5.4 Two-by-two Table of PET from the Reference Standard Data versus Data from the Ottawa Hospital Discharge Database

PET	Reference standard	
Ottawa hospital discharge database	Adjudicated as “ yes ”	Adjudicated as “ no ”
Coded as “ yes ”	37 (a) True positive	5 (b) False positive
Coded as “ no ”	66 (c) False negative	2758 (d) True negative

Sensitivity= $a/(a+c)$ =0.3592 or 35.92% (CI: 0.267-0.459)

Specificity= $d/(b+d)$ =0.9982 or 99.82% (CI: 995-0.999)

PPV= $a/(a+b)$ =0.8810 or 88.10% (CI: 0.74-0.94)

NPV= $d/(c+d)$ =0.9766 or 97.66% (CI: 0.97-0.98)

Disease incidence (in the OaK dataset according to reference standard) which called also reported incidence can be written as follows:

Disease incidence = $(a+c)/(a+b+c+d)$ =103/2866=0.0359 or 3.6%

Corrected Incidence of PET in the Ottawa hospital discharge database corrected for misclassification can be written as follows:

Corrected Incidence = (reported incidence x PPV)/sensitivity=0.036x0.88/0.36=0.088 or 8.8%

In Table 5.5, we presented the results comparing HDP results from the “reference standard” with presence or absence of HDP codes from the Ottawa hospital discharge database coded via ICD-10-CA in a traditional two-by-two table (Table 5.3).

Table 5.5 Two-by-two Table of HDP (PET and GH combined) from the Reference Standard Data versus data from the Ottawa Hospital Discharge Database

HDP	Reference standard	
Ottawa hospital discharge database	Adjudicated as “yes”	Adjudicated as “no”
Coded as “yes”	181 (a) True positive	18 (b) False positive
Coded as “no”	69 (c) False negative	2598 (d) True negative

Sensitivity= $a/a+c=0.724$ or 72.4% (CI: 0.67-0.78)

Specificity= $d/b+d=0.9931$ or 99.31% (CI: 989-0.995)

PPV= $a/a+b=0.9095$ or 90.95% (CI: 0.86-0.94)

NPV= $d/c+d=0.9741$ or 97.41% (CI: 0.97-0.98)

Disease incidence (in the OaK dataset according to reference standard) which called also reported incidence can be written as follows:

Disease incidence = $(a+c)/(a+b+c+d)=250/2866=0.872=8.72\%$

Corrected incidence (in the OaK dataset according to ICD-10-CA) can be written as follows:

Corrected incidence = $(\text{reported incidence} \times \text{PPV})/\text{sensitivity}=0.0872 \times 0.91/0.72=0.103$ or 10.3%

Sensitivity is a measure of the probability of correctly identifying a HDP/PET case through a presence of ICD-10-CA coding, and is synonymous with a true positive rate.

To calculate sensitivity we used the formula $a/a+c$ in the traditional two-by-two table where cell ‘a’ represents those HDP/PET positive according to the “reference standard”

database and in the Ottawa hospital discharge database (coded by ICD-10-CA) and cell ‘c’ represents those HDP/PET positive according to “reference standard” database and a HDP/PET negative according to the Ottawa hospital discharge database (coded by ICD-10-CA). In our study population, sensitivity was 0.3592 or 35.92% (CI: 0.27-0.46) for PET and 0.724 or 72.4% (CI: 0.67-0.78) for HDP.

We also calculated specificity, a measure of the probability of correctly identifying non-HDP/PET cases through hospital discharge database (coded by ICD-10-CA). To calculate specificity we used the formula $d/b+d$ in the traditional two-by-two table where cell ‘d’ represents those HDP/PET negative according to “reference standard” database and a HDP/PET negative according to the Ottawa hospital discharge database (coded by ICD-10-CA) and cell ‘b’ represents those HDP/PET negative according to “reference standard” database and a HDP/PET positive according to the Ottawa hospital discharge database (coded by ICD-10-CA). In our study population, specificity was 0.9982 or 99.82% (CI: 995-0.999) for PET and 0.9931 or 99.31% (CI: 989-0.995) for HDP.

Furthermore, we calculated the positive predictive value (PPV) defined as the probability that the person positive for the disease according to Ottawa hospital discharge database (coded by ICD-10-CA) is a true positive. To calculate PPV we used the formula $a/a+b$ and found to be 0.8810 or 88.10% (CI: 0.74-0.94) for PET and 0.9095 or 90.95% (CI: 0.86-0.94) for HDP.

Moreover, we calculated the negative predictive value (NPV) defined as the probability that the person negative for the disease according to hospital discharge database (coded by ICD-10-CA) is a true negative. To calculate NPV ($NPV = d/c+d$) and found 0.9766 or 97.66% (CI: 0.97-0.98) for PET and 0.9741 or 97.41% (CI: 0.97-0.98) for HDP.

5.5.3 Sample size re-calculation

Because the sensitivity of the Ottawa hospital discharge database for PET reporting (35.92%) was much lower than reported in the literature and used for calculating the sample size for our project, we have to re-calculate the sample size. The sensitivity used for sample size calculation was 70% and the lowest sensitivity used for sample size calculation in the reference material was 60% which is still higher than in our study; therefore, we have decided to use specificity instead of sensitivity for our sample size re-calculation as per the acceptable unit in the reference article.

Specificity was reported as 99.2%, prevalence of PET was 5%; [1, 2] by this way we were able to increase the minimal acceptable level of confidence interval with 0.95 probability from 0.55 to 0.92. [82] The number of cases to review was set as 77.

The corresponding number of controls is obtained from equation:

$$N_{\text{controls}} = N_{\text{cases}} [(1 - \text{Prev})/\text{Prev}] = 77 [(1 - 0.05)/0.05] = 1463$$

As a result, it was necessary to have 1463 non-PET subjects (controls) that make the total sample size of 1540 subjects. Our current sample size is, therefore, sufficient for our validation study.

With this established, we used the data from our study to see the influence of the calculated specificity and prevalence on the sample size. With specificity 95%, prevalence 0.15, and the number of cases as per reference table (Appendix 4), we have calculated the number of controls as required:

$$N_{\text{controls}} = N_{\text{cases}} [(1 - \text{Prev})/\text{Prev}] = 34 [(1 - 0.015)/0.015] = 2232$$

We have added the number of controls to the number of cases and calculated the total sample size which is equal 2266. This sample size is smaller than the sample size we

used in our study; therefore, we concluded that the sample size used in our study is sufficient to satisfy the power need.

5.5.4 Determinants of Misclassification

We performed the analysis to determine the influence of OaK study participants' factors on the misclassification of HDP and PET.

Our analysis was focused first on the comparison of 250/103 cases positive for HDP/PET, respectively, compiled in the reference standard with HDP/PET positive cases included in the Ottawa hospital discharge database (coded via ICD-10-CA).

The socio-demographic and clinical characteristics of the cases classified as HDP positive by the reference standard and the Ottawa hospital discharge database (correctly classified cases) were compared to those classified as HDP in the reference standard but not in the Ottawa hospital discharge database (misclassified cases) (Table 5.6).

We noted that the HDP misclassified cases tended to be younger (29.5 versus 30.5, $p < 0.0001$) and had lower BMI (26.7 versus 28.7, $p < 0.0001$); otherwise, there was no significant difference between the groups.

Furthermore, we compared the demographic and socioeconomic characteristics of the patients classified as PET cases in the reference standard and in the Ottawa hospital discharge database (correctly classified) to those classified as PET in the reference standard, and to those not classified as PET in the Ottawa hospital discharge database (misclassified cases) (Table 5.7).

Table 5.6 Demographic and pregnancy history characteristics for the correctly classified/misclassified HDP cases

Participant Factors		Not misclassified #/%	HDP (GH+PET) (misclassified cases) #/%	p value
Personal income	< 50 000\$	36 (20.34%)	12 (18.18%)	p=0.71
	> 50 000\$	141 (79.66%)	54 (81.82%)	
Mean age		30.50±5.33 (181)	29.49±4.62 (67)	p<0.0001
BMI		28.70±7.095 (170)	26.70±6.29 (68)	p<0.0001
Education	Secondary	26 (14.36%)	8 (11.59%)	p=0.56
	Post-secondary	155 (85.64%)	61 (88.41%)	
Repeat Caesarean Section	Yes	168 (92.82%)	62 (89.86%)	p=0.44
	No	13 (7.18%)	7 (10.14%)	
Hospital	Civic	28 (15.47%)	9 (13.04%)	p=0.62
	General	153 (84.53%)	60 (86.96%)	
EDD	Before 2006	124 (68.51%)	44 (65.67%)	p=0.67
	After 2006	57 (31.49%)	23 (34.33%)	
History of pregnancy complications*	Yes	148 (81.77%)	59 (85.51%)	p=0.48
	No	33 (18.23%)	10 (14.49%)	
History of PET	Yes	23 (12.78%)	6 (8.70%)	p=0.38
	No	157 (87.22%)	63 (91.30%)	
Parity	Nullipara	123 (67.96%)	44 (63.77%)	p=0.53
	Multipara	58 (32.04%)	25 (36.23%)	

*History of PET, placenta abruptio, or stillbirth in previous pregnancies

**Self-reported

Table 5.7 Demographic and Pregnancy History Characteristics for the Correctly Classified/Misclassified PET Cases

Participant Factors		Not misclassified #/%	PET misclassified, #/%	p value
Personal Income	< 50 000\$ > 50 000\$	6 (16.22%) 31 (83.78%)	18 (28.13%) 46 (71.88%)	p=0.175
Mean Age		30.62±5.46 (37)	29.9±5.63 (66)	p<0.0001
BMI		28.48±7.15 (35)	27.27±6.73 (60)	p<0.0001
Education	Secondary Post-secondary	4 (10.81%) 33 (89.19%)	12 (18.18%) 54 (81.82%)	p=0.32
Repeat Caesarean Section	Yes No	35 (94.59%) 2 (5.41%)	59 (89.39%) 7 (10.61%)	p=0.48 (Fisher)
Hospital	Civic General	6 (16.22%) 31 (84.21%)	4 (6.06%) 62 (93.94%)	p=0.16 (Fisher)
EDD	< 2006 > 2006	25 (67.57%) 12 (32.43%)	51 (77.27%) 15 (22.73%)	p=0.28
History of pregnancy complications*	Yes No	27 (72.971%) 10 (27.03%)	48 (72.63%) 18 (27.27%)	p=0.97
History of PET**	Yes No	4 (10.81%) 33 (89.19%)	9 (13.64%) 57 (86.36%)	p=0.77 (Fisher)
Severity***	Yes No	30 (81.08%) 7 (18.92%)	42 (63.94%) 24 (36.36%)	p=0.06
Parity	Nullipara Multipara	31 (83.78%) 6 (16.22%)	38 (57.58%) 26 (42.42%)	p=0.0067

*History of PET, placenta abruptio, or stillbirth in previous pregnancies

**Self-reported

*** According to SOGC classification

One significant result, as noted, was the observation that the misclassified PET cases tended to be younger (29.9 versus 30.62, $p<0.0001$) and had lower BMI (27.27 versus 28.48, $p<0.0001$) (Table 5.7).

The other significant finding was that the group of misclassified cases had a higher percentage of multiparous women than the correctly classified cases (42.42% versus 16.22%, $p=0.0067$). We attempted to uncover the reasons for this difference.

We calculated the incidence of PET and HDP based on parity in the reference standard database and in the Ottawa hospital discharge database and summarized the results in Table 5.8 and Table 5.9.

Table 5.8 Incidence of PET/HDP Based on Parity in the Reference Standard Data

Parity	Incidence of PET (%)	Incidence of HDP (%)
Nullipara	4.7	11.39
Multipara	2.4	5.93

Table 5.9 Incidence of PET/HDP Based on Parity in the Hospital Discharge Database

Parity	Incidence of PET (%)	Incidence of HDP (%)
Nullipara	2.39	9.07
Multipara	0.5	4.72

In this context, we found that the incidence of HDP in the Ottawa hospital discharge database was almost twice higher for nulliparous women than for multiparous women, but the incidence of PET was almost five-times higher among nulliparous women than among multiparous women. In the reference standard data, the incidence for both PET and HDP was almost twice higher among nulliparous women than multiparous women. From these results, we can draw the conclusion that the magnitude of misclassification in the Ottawa hospital discharge database is higher for multiparous women.

Another interesting observation was that the incidence of HDP was similar in both databases for both multiparous and nulliparous women (11.39% in the reference standard

database versus 9.07% in the hospital discharge database for multiparous women and 5.93% and 4.72% for nulliparous women). In contrast, the difference in the incidence of PET between databases was higher for both groups (nulliparous and multiparous).

The last finding was that there was an almost significant difference between misclassified and correctly classified cases based on the severity of PET (81.08% versus 63.94%. $p=0.06$). The percentage of less severe cases was twice as high among misclassified cases as among correctly classified cases (18.92% versus 36.36%).

We built a predictive model of PET misclassification based on the significance coefficient of the predictors. We hypothesized that factors such as age, BMI, parity, severity of PET, and history of chronic illnesses may be predictors of PET misclassification. Table 5.10 presents the data for Adjusted and Crude Odds Ratios Related to these Predictors. In this context, however, we found that only parity and severity of PET may serve as predictors of PET misclassification. The results show that multiparous women have an almost four times higher chance of being misclassified than nulliparous women (4.22, $p=0.0069$). The results also demonstrated that more severe PET cases have a three times lower chance of being misclassified than do less severe cases (0.33, $p=0.04$).

Table 5.10 Adjusted versus Crude Odds Ratios Related to Predictors of PET Misclassification

Predictor Variables	Crude OR with 95% CI	p-value	Adjusted OR with 95% CI	p-value
Age <30 (reference) ≥30	1.00 0.97 (0.43-2.19)	p=0.94	1.00 0.87 (0.36-2.1)	p=0.76
BMI <25 (reference) ≥25	1.00 1.07 (0.48-2.39)	p=0.86	1.00 1.43 (0.56-3.67)	p=0.45
Positive medical history Yes No (reference)	1.01 (0.41-2.5) 1.00	p=0.67	0.87 (0.32-2.39) 1.00	p=0.79
Para Nullipara (reference) Multipara	1.00 3.8 (1.40-10.36)	p=0.0089	1.00 4.22 (1.48-11.97)	p=0.0069
PET Severity Severe Not severe (reference)	0.48 (0.15-1.07) 1.00	p=0.06	0.33 (0.11-0.97) 1.00	p=0.04

5.5.6. Chart Review of the Misclassified PET Cases

We thoroughly reviewed all PET misclassified cases and compared them with the charts that were correctly coded in order to uncover the causes for misclassification. First, we searched the PET misclassified charts for the criteria for PET and found the following:

- BP readings (2 diastolic BP readings higher than 90 mmHg), along with positive urine protein dipstick reading ($\geq 2+$ and/or 24 hour urine protein $\geq 0.3\text{g/day}$) documented in 60% of the misclassified cases,
- HELLP syndrome (equivalent of PET) was documented in 3.6% of the cases, and
- either high symptomatic BP or abnormal lab work (high creatinine or liver enzymes) was mentioned in the charts of the remaining 36.4% of the cases (soft criteria for PET diagnosis).

We reviewed the charts, with focus on finding any documentation of PET diagnoses in the charts of misdiagnosed cases, and found that there was a clear documentation of the diagnosis for PET in the majority of these charts. This helped with the correct coding.

We also analyzed the charts of the misclassified cases for the consistency of diagnosis of PET recorded (for the purpose of this study, we did not separate mild PET cases from severe PET cases according to SOGC). In this context, we found that there was no consistency in the diagnoses recorded in the charts; there was a documented diagnosis of mild/severe PET, but we also encountered the following diagnoses: “high BP”, “severe GH with proteinuria”, “ $\uparrow$ BP with seizure”, “ $\uparrow$ BP with proteinuria”, “GH”, “hypertension”, “chronic hypertension”, “superimposed hypertension”, “chronic hypertension with superimposed PIH”, “mild to moderate PIH”, “PIH with preeclamptic toxemia”. The major finding was that terms “PIH” and “PET” were used interchangeably by clinicians.

Here are some examples of the comments found in misclassified cases:

Case #1: Documented more than two BP readings with $\text{dBP} \geq 90$ mmHg, 24 hour protein

urine protein was 1.05 g/day, +3 urine dipstick protein. Documented as “↑BP, proteinuria” on nursing notes, doctor’s notes and on delivery record as an indication for induction (not documented as GH or PET).

Case #2: Documented more than two symptomatic BP readings up to 180/120 mmHg, 24 hour urine protein was 0.65 g/day. Documented as “severe PET” on operative report, as “severe GH” on doctor’s progress notes, as “GH and severe PET” on delivery record.

Case #3: Documented more than two BP readings with dBP ≥ 90 mmHg, +2-3 urine dipstick protein, 24 hour protein urine protein was 1.04 g/day. Documented as “↑BP for 1 week” on nurse’s notes, “↑BP, proteinuria” on doctor’s notes, documented as “↑BP for 1 week” on the delivery record.

Case #4: Documented chronic hypertension with more than two BP readings with dBP ≥ 90 mmHg, +1-2 urine dipstick protein, 24 hour protein urine protein was not done, seizure. Documented as “↑BP with seizure” on doctor’s notes”, “PIH” on delivery notes, “↑BP, seizure during labor”, and “PIH” on doctor’s notes.

Case #5: Documented more than 2 BP readings with dBP ≥ 90 mmHg, +2 urine dipstick protein, 24 hour protein urine protein was 0.46 g/day. Documented as “↑BP from the beginning of pregnancy”, on nurse’s notes, as “superimposed PET and PIH” on doctor’s notes, and as “PIH” on the delivery record. .

The results of the chart review for PET cases according to the hospital discharge database but not according to the reference standard are presented below:

Case #1: Documented more than two BP readings with dBP ≥ 90 mmHg, negative urine dipstick protein, 24 hour protein urine protein was 0.26g/day, no adverse outcomes. Documented as “*PIH*” on doctor's notes, *and as “PIH” on delivery record.*

Case #2: Elevated BP. negative urine dipstick protein, no adverse outcomes. Documented as “*chronic*” throughout the chart, antenatal, as “*PIH*” on nursing and doctor's notes and “*PIH*” on the delivery record.

Case #3: Documented more than two BP readings with dBP ≥ 90 mmHg, +1 urine dipstick protein, 24 hour protein urine protein was 0.21 g/day, no adverse outcomes. Documented as “*elevated BP*” on operative report, nursing notes and as “*PIH*” on delivery record.

Case #4: Documented more than two symptomatic BP readings with dBP ≥ 90 mmHg, negative urine dipstick protein, 24 hour protein was 0.22 g/day, no adverse outcomes. Documented as “*PIH*” on doctor's notes, “ $\uparrow$ *BP*” on nursing notes, “*elevated BP/PIH*” documented on delivery record.

Case #5: Elevated BP. negative urine dipstick protein, no adverse outcomes. Documented as chronic throughout the chart, antenatal, on the delivery record where it's documented as "*PIH*", "*essential hypertension*" and "*chronic hypertension*"

After further review, it has been established that these cases have not fulfilled criteria to be classified as PET in accordance with SOGC Guideline.

We identified three different kinds of errors:

1. PET pregnancies coded with incorrect ICD-10-CA code;
2. Pregnancies affected by PET coded as normal;
3. Pregnancies adjudicated as other than PET coded as PET

Among misclassified cases, 56 PET cases according to the reference standard (78.88%) were coded with incorrect ICD-10-CA code in the Ottawa hospital discharge database: 3 (4.23%) were coded as "O10.0 pre-existing essential hypertension", 53 (74.65%) as "O13.0 Gestational [pregnancy-induced] hypertension without significant proteinuria), and 10 PET cases (14.08%) were not coded as hypertension of any kind (normal according to ICD-10-CA). Also there were 5 (7.04%) not PET cases according to the reference standard (3 cases adjudicated as PIH and 2 as CH) but coded as "O14.0 Gestational [pregnancy-induced] hypertension with significant proteinuria" in the Ottawa hospital discharge database (Table 5.11).

Table 5.11 Frequencies of Different Kinds of Errors in the Ottawa Hospital Discharge Database

Errors	Frequency (#)	Percentage (%)
PET pregnancies coded with incorrect ICD-10-CA code	56	78.88
Pregnancies affected by PET coded as normal	10	14.08
Not PET pregnancies coded as PET	5	7.04

5.6 Discussion

5.6.1 Validation studies to improve PET reporting in administrative databases

PET is a complication of pregnancy that can affect both the mother and the infant with incidence 3.0-10%. [2, 3] It is expected that the difference in incidence is partially due to misclassification of this condition. [81] In our cohort study (Oak Birth Cohort) the incidence of PET was 3.6% which was different from the incidence of PET according to the hospital discharge database which was 1.47 %. The incidence of HDP in our study population was 8.72% but 5.43% in the hospital discharge database. Our results have demonstrated that HDP, especially PET, were underreported in the hospital discharge database.

Validation studies focusing on reporting of HDP in hospital discharge databases have been conducted in Europe, Australia, and North America and have shown similar results: HDP were underreported in hospital discharge databases; the only difference among countries was the magnitude of underreporting. [81]

To our knowledge this is the first Canadian hospital discharge database validation study conducted in the province of Ontario and that used adjudicated data from an observational study as the “reference standard” instead of chart review. Another

Canadian study that validated hospital discharge database diagnoses by using the Nova Scotia Atlee Perinatal Database as a reference standard, showed similar results: low sensitivity for PET (although higher than in our study) with similar specificity (99.5%). [79] The results of our study were similar to those of other validation studies. The magnitude of the underreporting was much higher in our study than in other studies with similar objectives: Roberts et al. [64] reported sensitivity of PET reporting ranging from 71.0 to 84.7% and specificity 96-99.2% (18% and 99.8% respectively in our study) and for any pregnancy hypertension 63.3% to 68.2% (23% in our study) with PPV for PET 31.6-66.7% (98% in our study). In another Canadian study, sensitivity for gestational hypertension with proteinuria (severe GH or HELLP) was 75.2% and for any gestational hypertension was 87.9% with specificity 99.5% and 99.6% respectively (PPV was not presented). [79] Klemmensen et al. reported low sensitivity values for both PET and HDP ranging from 43.64 to 48.88% with high sensitivity and PPVs reported in the Danish National Patient Registry administrative database (coded by ICD-10). Overall 64% of PET cases were misclassified in hospital discharge database. [78]

We compared the incidence of PET and HDP from the hospital discharge abstracts with the reference standard and found that the incidence of both conditions reported in the hospital discharge database was twice lower than in the reference-standard. The results have shown that both conditions were severely underreported in the Ottawa hospital discharge database. Roberts et al. reported that the incidence of PET and HDP in hospital discharge databases varied greatly: as low as 1.4% for PET and 6.0% for HDP in Alberta Canada to as high as 4.0% for PET (Norway) and 9.1% for HDP (Australia). [81] The incidence of PET and HDP reported in hospital discharge database in the Alberta Health

region was as low as 0.6% and 6.3% respectively. [80] The reported data on Canadian incidence of PET/HDP presents the lowest incidence among all countries studied. We have not found evidence of any study conducted in the province of Alberta to validate reporting of PET and HDP in general. The only Canadian study conducted to validate reporting of pregnancy complications did not report the incidence of PET and HDP. [79] Because there was no study conducted in the province of Alberta to validate the reporting of HDP, the question remains whether the data presented is reliable or has been influenced by a high magnitude of underreporting.

The data on HDP/PET incidence was used to compare the incidence of these conditions in Canada with that in other countries such as Australia, Denmark, Scotland, Norway, and Sweden, as well as one hospital in the USA. [81] The incidence of PET in Canada was the lowest among these countries. Authors extrapolated this data to represent the incidence of this condition in the whole country; they hypothesized that because of Canadian population's homogeneity the incidence of PET in Canada should be similar to its incidence in Alberta. The results of our study have shown that the incidence of PET in hospital discharge database of the province of Ontario is, in fact, similar to that in Alberta. Because there has been no hospital discharge database validation study on PET diagnosis conducted in either of these provinces, we cannot be sure that incidence of PET derived from hospital discharge databases represents the true incidence of PET. Further studies should be conducted to validate the use of hospital discharge databases for calculation of PET incidence. The PET definition used by clinicians in different hospitals and provinces should also be taken into consideration.

The importance of these findings cannot be overstated. If prevalence of this condition is used to evaluate the epidemiological trends of this condition, it might show the reduction of this condition in certain areas or over time which could be misleading; it could have a negative effect on the distribution of hospital funding, as well as policy implications.

Majority of the epidemiology scholars must accept the fact that there will be some degree of PET misclassification in hospital discharge data based on the data we have presented, and that is acceptable as long as the misclassification is not at excessive level. That is why, in our opinion, validation studies are very important. They would allow quantification of the magnitude of PET misclassification. In the literature, however, we have not found the range for sensitivity of hospital discharge database in PET ascertaining which would be considered acceptable to decide whether a database is accurate enough for this purpose. In this regard, we would expect that sensitivity which is equal or greater than 70% would be considered acceptable. Another way to use data from hospital discharge databases is by combining data on PET and PIH/GH because it is known that these two diagnoses are often used interchangeably.

5.6.2 The Effect of Patient Characteristics on PET Misclassification

In our study we attempted to explore the effect of basic patient identifiers on PET misclassification. The results have shown that the misclassified cases tended to be younger (29.9 versus 30.62, $p < 0.0001$) and with lower BMI (27.47 versus 28.48, $p < 0.0001$). Even though the difference in mean age and BMI between groups was statistically significant, we do not think that it was clinically significant.

We found that multiparous women diagnosed with PET had a significantly ($p=0.0067$) higher chance to be misclassified than nulliparous women with PET. We hypothesized that there could be two explanations for this finding. First, nulliparity is a known PET risk factor; the incidence of PET is higher in nulliparous women. [7, 70] We calculated the incidence of PET in a group of nulliparous women versus a multiparous group and found similar results (4.5% versus 2.4% according to the adjudicated data and 2.39% versus 0.5% according to hospital discharge database). A second possible explanation is implicit in the first explanation: nulliparous women could have gotten more attention from the health care professionals during their antenatal period and during labor because of the fact that they have a higher chance of developing pregnancy complications such as PET. The data show that the incidence of PET reported by the hospital discharge database was twice less than the one from the reference-standard in the group of nulliparous women, but in the group of multiparous women, the incidence of PET reported by hospital discharge database was almost 5 times less than the one from the reference-standard.

5.6.3 Classification of Errors and Potential Causes of Misclassification

We developed a classification of errors and found that majority of the cases belong to the category “pregnancies affected by PET coded with incorrect ICD-10-CA code” (78.88%) followed by cases belong to the category “pregnancies affected by PET coded as normal” (14.08%). There were 7.04% pregnancies cases in the category “not affected by PET coded as PET”. Our results are similar to the results of Thornton who reported an overall error rate of 56-64% for HDP, depending on the database studied. [73] We planned to use

different classification of coding errors based on incorrect versus correct diagnosis recorded in the charts but because of the diversity of the diagnoses used by physicians, we abandoned this idea.

Additionally we explored the potential causes of misclassification. The most significant finding was the diversity of diagnoses recorded in the medical charts. We think that it could have contributed to PET misclassification due to the fact that two diagnoses PIH and PET, are often used interchangeably. [15] Larsen et al. reported similar results in the study that evaluated the accuracy of diagnoses of venous thromboembolism during pregnancy and postpartum in hospital discharge summary. The common error that occurred as a result of interchange between the diagnosis of deep venous thrombosis and superficial thrombophlebitis was seen in 8.2%. [43]

According to some authors, [75] correct reporting depends on the severity of the PET. The probability of correctly coding severe PET is higher than that for mild PET or GH. Our findings were congruent with the author's; we found that there was an almost significant difference between misclassified and correctly classified cases based on PET severity ($p=0.06$). The percentage of less severe cases was twice higher among misclassified cases than correctly classified cases.

We, furthermore, discovered that HDP classification used by clinicians differ from the one in the ICD-10-CA, most probably because clinicians are not aware of ICD-10-CA definitions. We realize that the main goal of the clinicians is to provide the best possible care for their patients but the purpose of SOGC Guideline is to facilitate communication among caregivers, and to create meaningful groups with different prognoses, considerations for surveillance, and/or outcomes. [2] This Guideline was developed by

practicing physicians who provide the obstetrical care to guide the fellow physicians to proper management of the patients with HDP. They advise that “the term PIH (pregnancy-induced hypertension) should be abandoned, as “its meaning in clinical practice is unclear” and we hope that introduction of this Guideline into clinical practice will not only help provide standardized care for women with HDP but also educate clinicians on proper terminology for HDP and, potentially, improve coding of this condition as well.

Up to now, the identification of this serious pregnancy complication depends on the experience and skills of a coder, who may not necessarily have enough medical background and knowledge in obstetrics to properly identify and code it. Knowing that these limitations might have complicated the coding, we obtained the training material used by coders to facilitate their work, and we discovered that it did not even cover hypertensive disorders of pregnancy.

We think that discharge summary written by physicians involved in the care of the patient (required for all patients staying greater than 2 days) could have helped coders to identify and properly code HDP but, unfortunately, we found that it was not available in the majority of the charts. Patients who stayed less than 2 days (majority of women with no complications do not stay in the hospital for more than 48 hours) do not require a discharge summary.

5.6.4 Strengths and limitations

The strength of our study is that it is a validation study, and all the cases were adjudicated by a panel of clinicians experienced in perinatology.

The study has limitations. One of them is that the chart review was done retrospectively; therefore, we have no knowledge whether all information (clinical, discharge summaries and lab data) available to us was available to medical coders who worked in real time format. However, according to SOGC, it is a fact, that a woman with HDP may meet criteria for more than one condition, and that evolution may occur over time (i.e., a woman may be admitted with GH which would progress into PET over a period of inpatient admission). A final diagnosis of the type of HDP, therefore, should be retrospective, following the postpartum period. [2]

Another limitation is lack of knowledge whether proper procedure for BP was used. According to the SOGC Guideline, [2] diagnosis of PET is based on diagnosis of hypertension and proteinuria. The Guideline recommends that hypertension in pregnancy should be defined as a diastolic BP of ≥ 90 mmHg, based on the average of at least two measurements, taken using the same arm. There are recommendations on proper procedure for BP measurement as well such as “BP should be measured with the woman in the sitting position with the arm at the level of the heart”, “an appropriately sized cuff should be used”, “Korotkoff phase V should be used to designate diastolic BP”, however, it is unknown whether these recommendations were followed by all obstetrical care providers. The BP measurement procedure is usually not documented in the chart and, therefore, retrospective chart review does not provide information on whether diagnosis of hypertension in pregnancy was done in accordance with the SOGC Guideline. On the other hand, this limitation is not innate to our study only but rather inherent in any study based on data from hospital discharge database.

Our study was done only in one hospital in the province of Ontario and, therefore, may be not representative of the coding process in other hospitals of the province or Canada, in general. However, the incidence of PET obtained from the Ottawa hospital discharge database was similar to the incidence of PET according to Niday Perinatal Database (Ontario Provincial database) (1.47% and 1.3% respectively) [87] and to the incidence of PET in one of the hospital of the province of Alberta according to hospital discharge database (1.47% and 1.4% respectively). [80] Furthermore we compared the mean age of the OaK study population with the mean age for all births in Ontario which was 30.0 ± 5.5 years. The mean age of mothers in the OaK study was 30.35 ± 5.06 . According to the Ontario Perinatal Surveillance System Report 2008, [87] over 80% of women in Ontario received care from a physician over the course of their pregnancy, while ~5% of women only use the services of midwives. Almost all women in Ontario give birth in hospital, with only 1.6% of women give birth at home. We believe that based on these facts, recruitment from the hospitals during antenatal visits, provided a representative sample of women giving birth in the whole province of Ontario.

Finally, we used the adjudicated data from the OaK study as the “reference standard”. PET definition used was adjudication was more inclusive than restrictive, and this may suggest the necessity of some caution in the interpretation of results. There are still debates among PET scholars about definition of PET. According to some scholars, the definition should be simple, clinically useful and allow prediction of negative pregnancy outcomes. [14, 16] The problem is that clinicians are looking for a PET definition that would allow predicting poor obstetric outcomes, but epidemiology scholars are looking for a definition that would allow using PET definition to determine the PET trends and

compare them among different hospitals in the same country or among different countries.

5.6.5 Proposed methods to improve PET reporting

We believe that PET reporting can be improved. One of the solutions is to use different sources of administrative databases to match cases; another solution is to educate physicians about the importance of proper ICD-10-CA coding and classification used for coding of hypertensive disorders of pregnancy; and the coders about importance of communication between medical coders and physicians to achieve better results in coding this serious complication of pregnancy.

We believe that the PET definition should be chosen carefully based on the objectives of the study. For epidemiologic research, the restrictive definition is more suitable because it has fewer criteria open to interpretation. It also allows comparison of results with other studies, given the same PET definition used.

Another way to improve the PET reporting is through creation of a disease (PET) specific database.

To conclude, we would like to consider two very important things. First, that evidence-based principle of practice is a paramount of modern medicine but it relies significantly on the accuracy of data obtained from different sources. Second, it is a fact that the administrative databases are becoming one of the main sources for the data used in research; the Nova Scotia Atlee Perinatal Database and the Ontario Perinatal Surveillance System are two examples of such databases. Clinicians use these databases to research the epidemiology of diseases and develop guidelines for improving the quality of care for

the patients; therefore, the importance of the accuracy and data quality in these databases is difficult to overestimate.

5.7 Conclusion

PET was severely underreported in the hospital discharge database of the study population. There is perhaps a huge gap between information likely available to clinicians and to medical coders. The findings demonstrated that clinicians are not aware of either the importance of proper documentation for the correct diagnosis or the proper coding procedure and ICD-10-CA codes used to code hypertensive disorders of pregnancy. The medical coders do not have the necessary tools to help them identifying and correctly coding hypertensive disorders of pregnancy.

Chapter 6 Project 3

Evaluation of the Effect Outcome Misclassification on the Results

6.1 Aim

The aim of this project is to estimate how misclassifications of PET in administrative databases may result in different conclusions.

6.2 Hypothesis

The estimate of association between PET and some of the risk factors expressed in ORs depends on the database the outcome was obtained from.

6.3 Background

The development of information technology during the last decades has resulted in electronic storage of an increasing amount of data. This data increasingly used not only in epidemiologic studies but for health care planning and reimbursement systems as well; therefore it is of great importance that administrative databases have a high degree of reliability and validity. The quality of routinely collected data for administrative purposes has been questioned by some researchers. [73-76] In our study we have shown that the Ottawa hospital discharge database coded via the ICD-10-CA system underreported one

of the serious complications of pregnancy, PET. In this project we will explore the effect of such misclassification on the parameter estimate of risk factor outcome association.

6.3.1 PET Risk Factors

Even though PET is the leading cause of maternal and fetal mortality and morbidity in developed countries, the etiology of this disease remains mysterious. [1] On the other hand, a lot of research has been done, to date, to establish the risk factors of PET in order to develop preventive strategies for this condition. Below, we present the most common risk factors of PET.

6.3.1.1 Age

Most of the results of the studies of PET risk factors agreed that PET is a disease of extreme age; extremely young (teenage) or relatively old (over 35) women are at an increased risk of developing PET. [70] Data showed that the risk of PET increased by 30% for every additional year after age 34. [88]

6.3.1.2 Parity

Women in their first pregnancy are at an increased risk of PET. A few studies have shown that nulliparity almost tripled the risk for PET. [89-91]

6.3.1.3 PET in previous pregnancy

Previous pregnancies appear to have a protective effect. But women with a history of PET in previous pregnancies are at an increased risk; women who had PET in a first pregnancy have seven times higher risk of developing PET in a second pregnancy.

6.3.1.4 Multiple pregnancy

The studies showed that the risk of developing PET nearly tripled if a woman carried a twin pregnancy in comparison with a woman who carried a singleton pregnancy. [94, 95]

6.3.1.5 Pre-existing diabetes mellitus

Research indicated that some chronic diseases were associated with PET. The likelihood for PET development almost quadrupled for women who suffered from pre-pregnancy diagnosed diabetes. [96, 97]

6.3.1.6 Chronic hypertension

Women with pre-existing hypertension were at higher risk of developing PET than those without this condition. [98]

6.3.1.7 Body mass index

Majority of the research on prepregnancy BMI and PET has shown that women with increased BMI were at higher risk of developing PET. [91, 97, 99]

6.3.1.8 Smoking

A few studies reported that smoking had a protective effect on PET development; women with a smoking history demonstrated lower probability of developing PET. [100, 101, 102] One study reported that smoking did not have a protective effect on women who smoked but quit before becoming pregnant. [103]

6.3.1.9 Race

It has been reported that some ethnic groups such as African-American and Hispanic women had a higher incidence of PET when compared to White women. [104]

6.3.2 Introduction of Concept Differential versus Non-Differential Disease

Misclassification

6.3.2.1 Definitions

Classification error that depends on the actual values of other variables is called differential misclassification. Classification error that does not depend on the actual values of other variables is called non-differential misclassification. Classification error that depends on the errors in measuring or classifying other variables is called dependent error; otherwise the error is called independent or nondependent error. The bias caused by differential misclassification can either exaggerate or underestimate an effect. Non-differential disease misclassification occurs when the proportion of subjects misclassified on disease does not depend on the status of the subject with respect to other variables in the analysis, including exposure. Popular view is that bias introduced by non-differential misclassification of a binary disease is predictable in direction, toward the null value to be exact, considering all other conditions are met. [56] First, published rules assume that

the misclassification probabilities are exactly non-differential; small violations of this assumption can produce substantial bias away from the null. Second, the exposure misclassification errors are assumed to be independent of errors in other variables in the analysis. Third, further conditions are required to guarantee bias towards the null when the exposure is polytomous (>2 levels). Fourth, the rules assume absence of interactions with other sources of systematic error, such as selection bias and confounding. In practice it is difficult to guarantee that all these conditions are satisfied and common practices often lead to violations of the assumptions. [105]

In our project we will attempt to demonstrate the direction and magnitude of the bias introduced by non-differential misclassification. We do not expect misclassification being differential given that the outcome occurred after the exposure and outcome ascertainment was blinded to the exposure status. [56]

6.4 Methods

6.4.1 Study Design

A retrospective cohort study design

6.4.2 Population

- **OaK Birth Cohort**

As described in preceding chapters, the OaK Birth Cohort Study recruited pregnant women in the Ottawa Hospital and Kingston General Hospital and followed until birth. Between October 2002 and April 2008 the OaK Birth Cohort 8075 mothers had been recruited. The aim of the study was to ascertain the association between folate intake

during pregnancy and adverse placenta-mediated pregnancy outcomes including intrauterine growth restriction (IUGR), PET, placenta abruptio (PA), and fetal death.

To be included in the study, women had to be between 12 and 20 weeks of gestation, with a viable fetus, and an expected singleton or twin pregnancy. The following data was obtained for each participant of OaK study:

- Demographic data (date of birth, weight, height, income, education, ethnicity);
- Past pregnancy history (history of PET, PA, stillbirth, or gestational diabetes in previous pregnancies);
- Pre-existing medical conditions (history of chronic hypertension, diabetes, or other chronic conditions);
- Social History (smoking, alcohol consumption during pregnancy);
- Delivery information (mode of delivery, date of birth of the baby, gestational age at birth, status of the baby at birth, weight of the baby and other variables);
- Pregnancy complications (PA, HDP/PET)

This demographic and health information was collected by using a CRF administered by the obstetrical research nurses. Although demographic, past pregnancy history, social history, pre-existing medical conditions data were self-reported by the study participants, the information was also verified against information from the participant's medical records to ensure the accuracy of the data. Pregnancy outcomes data were later collected from hospital charts after delivery to complete the CRF for each participant. Infant data, such as sex, weight, and other variables were also collected at birth. 5277 women were

recruited from the Ottawa Hospital and 2808 from the Kingston General Hospital. Recruitment statistics at the Ottawa Hospital is as following:

5277 women were recruited from Ottawa Hospital and 2808 from Kingston General Hospital. Majority of the participants delivered at Ottawa Hospital General campus, followed by Ottawa Hospital Civic campus. Minority of the participants delivered in other Ottawa hospitals (Montfort, Queensway Carleton) and elsewhere (other provinces). The exact distribution of participants based on delivery hospital was unknown.

Adjudicated data from OaK Study was considered as a “reference standard” for this project.

- **Administrative database**

The Ottawa hospital discharge database was considered as administrative database for this project.

6.4.3 Statistical Analysis

6.4.3.1 Dependent Variables

Dependent variable is PET.

6.4.3.2 Independent Variables

Independent variables were included in the regression models were maternal age, ethnic background, educational level, parity, previous preeclampsia, chronic hypertension, diabetes, pre-pregnancy BMI, household income, and cigarette smoking during pregnancy.

6.4.3.3 Descriptive Statistics

Descriptive statistics were calculated for all socio-demographic and clinical characteristics using means and standard deviations for continuous variables and frequency counts and percentages for categorical factors. Comparisons between groups were performed using cross-tabulations and either Fishers' exact or chi-square tests of significance.

6.4.3.4 Differences between Groups

The Student t-test was used for normally distributed continuous variables.

The Mann-Whitney test was used for non-normally distributed variables.

Chi -square test or Fisher's exact test was used for categorical variables.

6.4.3.5 Calculation of Risk

Odds ratio (OR) was used as an approximation of the relative risk given low incidence of the outcome (PET). An OR approximates how much more likely (or unlikely) the outcome is for those with the condition present ($X=1$) than among those with its absent ($X=0$).

6.4.3.6 Correlation

Spearman's rank correlation was used for a correlation assessment. A correlation coefficient varies from +1 to -1. If it is zero the variables are not related. If it is positive, these are positively correlated: one increases when the other increases. If it is negative,

these are negatively correlated: one increases when the other decreases and vice versa.
[106]

6.4.3.7 Logistic Regression Analysis

As a first step, univariable analysis was employed to evaluate individual variables as potentially significant predictors of PET.

Then, multivariable logistic regression analysis was employed to assess the contribution of each of these variables to our dichotomous outcomes. The significant level for entering a variable into the model was set as 0.05 and as 0.10 for a variable to remain in the model. [107]

Hosmer and Lemeshow Goodness-of-Fit test, chi-square goodness-of-fit test was used to test whether the model's estimates fit the data at an acceptable level or not. [108]

All analyses were performed using SAS (version 9.2 SAS Institute Inc., Cary, NC).

6.5 Results

6.5.1 Maternal Demographic Characteristics

A total of 2887 women were included in the analysis. Socio-demographic characteristics of the cohort are presented in Table 6.1. The majority of the study participants were Caucasian with high socioeconomic status, high education level, normal weight, and non-smoking.

Next, we evaluated whether there was a difference in socio-demographic characteristics between the group of women who developed PET and those who did not in accordance with the reference standard data. Women who developed PET were more likely to be

younger, to belong to Black or Asian ethnic groups, to have lower education level and household income, to be underweight or overweight [109]; and to smoke cigarettes during pregnancy more than women who did not develop PET. As for known risk factors, our results showed that women who developed PET were nulliparous, carried twin pregnancy, more often had pre-existing hypertension, diabetes type 1 (not type 2) and a history of PET in previous pregnancies. Refer to Table 6.1 for the detailed data.

Table 6.1 Socio-demographic Characteristics of Women according to PET Status

Socio-Demographic Factors	Overall	No PET 2763	PET 103 (3.6%)	Chi² (F-Fisher)
Age (Mean)	2866 30.91±4.83	2763 30.95±4.8	103 30.16±5.55	<.0001
Age < 25 25-29 30-35 >35	270 (9.42%) 775 (27.04%) 1150 (40.13%) 671 (23.41%)	254 (9.19%) 746 (27.00%) 1115 (40.35%) 648 (23.45%)	16 (15.53%) 29 (28.16%) 35 (33.98%) 23 (22.33%)	0.16
Ethnicity Aboriginal Black Asian Caucasian Other	28 (0.98%) 80 (2.81%) 199 (6.98%) 2142 (75.11%) 403 (14.13%)	26 (0.95%) 74 (2.69%) 191 (6.95%) 2070 (75.30%) 388 (14.11%)	2 (1.94%) 6 (5.83%) 8 (7.77%) 72 (69.90%) 14 (14.56%)	p=0.20 (F)
Family Income < 25 000\$ 25000-49 999\$ 50 000-79 999\$ ≥80 000\$	162 (6.06%) 312 (11.66%) 648 (24.22%) 1553 (58.06%)	155 (5.79%) 295 (11.46%) 619 (24.05%) 1514 (58.47%)	7 (6.93%) 17 (16.83%) 29 (28.71%) 48 (47.52%)	0.14
Income <50 000\$ >50 000\$	474 (17.72%) 2201 (82.28%)	450 (17.48%) 2124 (82.52%)	24 (23.76%) 77 (76.24%)	0.10
BMI	2796 24.58±5.38	2701 24.47±5.29	95 27.71±6.87	<.0001
BMI <18.5 (underweight) 18.5-24 (normal weight) 25-29 (overweight) ≥30 (obese)	109 (3.90%) 1740 (62.23%) 570 (20.39%) 377 (13.48%)	106 (3.92%) 1702 (63.01%) 546 (20.21%) 347 (12.85%)	3 (3.16%) 38 (40.00%) 24 (25.26%) 30 (31.58%)	<.0001
Education				

Grade School	20 (0.70%)	19 (0.69%)	1 (0.97%)	0.24 (F)
High School	294 (10.18%)	276 (10.01%)	15 (14.56%)	
Graduate	252 (8.81%)	241 (8.74%)	11 (10.68%)	
completed not completed	2296 (80.31%)	2220 (80.55%)	76 (73.79%)	
Para	1466 (51.17%)	1497 (50.58%)	69 (66.99%)	0.001
Nullipara	1399 (48.83%)	1365 (49.42%)	34 (33.01%)	
Multipara				

Table 6.1 Socio-demographic Characteristics of Women according to PET Status (cont.)

	Overall	No PET 2763 (96.4%)	PET 103 (3.6%)	Chi² (F-Fisher)
Smoking				
Yes	182 (6.35%)	174 (6.30%)	8 (7.77 %)	0.54
No	2679 (93.87%)	2587 (93.70%)	95 (92.23%)	
Education				
Secondary	311 (10.88%)	295 (10.70%)	16 (15.53%)	0.12
Post-secondary	2548 (89.12%)	2461 (89.30%)	87 (84.47%)	
Chronic hypertension (self-report)				
Yes	45 (1.57%)	32 (1.16%)	13 (12.62%)	<.0001
No	2821 (98.43)	2731(98.84%)	90 (87.38%)	
Chronic hypertension (according to the chart)				
Yes	53 (1.86%)	37 (1.35%)	16 (15.84%)	<.0001
No	2795 (98.14%)	2710 (98.65%)	85 (94.16%)	
Diabetes 1				
Yes	33 (1.15%)	26 (0.94%)	7 (7.00%)	<.0001
No	2829 (98.85%)	2736 (99.06%)	93 (93.00%)	
Diabetes 2				
Yes	36 (1.26%)	33 (1.20%)	3 (2.97%)	0.13 (F)
No	2826 (98.74%)	2728 (98.80%)	98 (97.03%)	
Twin pregnancy				
Yes	51 (1.79%)	42 (1.53%)	9 (8.82%)	<.0001
No	2803 (32%)	2710 (98.47%)	93 (91.18%)	
Hx of PET (self-report)				
Yes	94 (3.25%)	80 (2.90%)	13 (12.62%)	<.0001
No	2769 (96.75%)	2679 (97.08%)	90 (87.38%)	

Once this has been established, we evaluated the association between potential risk factors and the development of PET.

Predictor variables were tested one by one to obtain crude odds ratios (ORs) with 95% CIs. After testing for collinearity (using Spearman's correlation coefficient) and interaction, all covariates tested in the univariate analysis were included in the final model to obtain adjusted ORs and 95% CIs. For the dependent variable of interest adjusted ORs and corresponding CIs are presented, see Table 6.2 for results.

Table 6.2 Adjusted versus Crude Odds Ratios Related to Predictors of PET

Predictor Variables	Crude OR with 95% CI	p-value	Adjusted OR with 95% CI	p-value
Ethnicity				
Black	2.33 (0.98-5.53)	p=0.09	2.68 (0.91-7.89)	p=0.07
Asian	1.2 (0.57 -2.53)	p=0.74	2.44 (1.09-5.48)	p=0.03
Caucasian (reference)	1.00		1.00	
Other	1.11 (0.63-1.96)	p=0.46	0.87 (0.44-1.69)	p=0.67
BMI				
<18.5 (underweight) reference	1.00		1.00	
18.5-24 (normal weight)	0.79 (0.24-2.60)	p<.005	0.93 (0.27-3.21)	p=0.67
25-29 (overweight)	1.55 (0.46-5.25)	p=0.62	2.04 (0.56-7.35)	p=0.27
≥30 (obese)	3.05 (0.91-10.217)	p<.0003	3.31(0.92-11.91)	p=0.067
Age				
< 25	2.00 (1.09-3.68)	p=0.03	1.96 (0.85-4.53)	p=0.11
25-29	1.24 (0.75-2.04)	p=0.79	1.17 (0.65-2.09)	
30-35 reference	1.00		1.00	p=0.60
>35	1.13 (0.66-1.93)	p=0.46	1.31 (0.71-2.43)	p=0.38
Family Income				
<25000\$	1.42 (0.63-3.19)	p=0.96	2.55 (0.75-8.53)	p=0.13
25000-49 999\$	1.81 (1.02-3.19)	p=0.23	2.51 (0.75-8.33)	p=0.13
50 000-79 999\$	1.47 (0.92-2.35)	p=0.77	1.83 (0.54-6.27)	p=0.33
≥80 000\$ (reference)	1.00		1.00	
Education				
High school or below (reference)	1.00		1.00	
Graduate (including not complete)	0.65 (0.38-1.13)	p=0.12	1.1(0.49-2.43)	p=0.81
Smoking				
Yes	1.25 (0.60-2.62)	p=0.55	1.31 (0.53-3.24)	p=0.56
No (reference)	1.00		1.00	
Chronic hypertension				

Yes No (reference)	12.32 (6.26-24.28) 1.00	p<.0001	6.33 (2.79-14.36) 1.00	p <.0001
Diabetes 1 Yes No (reference)	7.93 (7.93-18.72) 1.00	p<.0001	8.94 (3.29-24.31) 1.00	p<.0001
Diabetes 2 Yes No (reference)	2.53 (0.76-8.39) 1.00	p =0.13	0.42 (0.05-3.53) 1.00	p=0.42

Table 6.2 Adjusted versus Crude Odds Ratios Related to Predictors of PET (cont.)

Predictor Variables	Crude OR with 95% CI	p-value	Adjusted OR with 95% CI	p-value
Hx of PET Yes No (reference)	4.84 (2.60-9.02) 1.00	p<.0001	5.71 (2.48-13.18) 1.00	p<.0001
Para Nullipara Multipara (reference)	1.98 (1.5-2.50) 1.00	p<0.001	2.83 (1.64-4.89) 1.00	p<.0001
Twin pregnancy Yes No (reference)	6.24 (2.95-13.21) 1.00	p<0.0001	6.02 (2.33-15.75) 1.00	p=0.002

In our study, the following factors were statistically significantly associated with PET:

- Asian race (adjusted ORs 2.44; Caucasian as a reference);
- Chronic hypertension (adjusted OR 6.3);
- Diabetes type 1 (adjusted OR 8.94);
- History of PET in previous pregnancy (adjusted OR 5.71);
- Nulliparity (adjusted OR 2.83);
- Carrying twin pregnancy (adjusted OR 6.02).

To assess the effect of the outcome misclassification on the parameter estimate, the logistic regression analysis was repeated using PET incidence data obtained from the hospital discharge database via ICD-10-CA coding. The ORs from this model were compared with the ORs obtained through the logistic regression analysis using the

reference standard data for the incidence of PET. The results are presented in table 6.3. In general, those factors that acted as risk factors for PET development (Asian race, chronic hypertension; diabetes type 1, history of PET in previous pregnancy, nulliparity, and twin pregnancy) continued to exhibit the same qualities. Even though the protective effect of smoking on PET has been reported by some researchers, our results do not support this effect.

In turn, we examined the extent to which the estimate of the association between PET and some of the risk factors expressed in ORs changed depending on whether the incidence of PET was based on the reference standard data or the data from the hospital discharge database. The selection of the risk factors was based on the following criteria:

- Known risk factors (from the literature and based on univariate analysis): age, BMI, parity, history of PET in previous pregnancy, pre-pregnancy hypertension and pre-pregnancy diabetes type 1;
- Controversial risk factors (e.g. smoking; some studies showed protective effect; some did not)

The results from table 6.3 demonstrate that those variables considered strong predictors of PET did not depend on the method of ascertaining PET incidence. All these variables (i.e., parity, history of PET, pre-existing hypertension and diabetes type I) show a strong association with PET. For diabetes type 1, loss in significance was observed, most probably, due to the small sample size. Even a weak predictor like smoking exhibits the same tendency. Another interesting observation concerns the effect of twin pregnancy on PET incidence. When PET incidence was ascertained through the reference standard data, it showed a significant positive association with

the development of PET, however, when it was ascertained through the hospital discharge database, the significance of the association was lost ($p<.0001$ versus $p=0.13$).

Table 6.3 Adjusted ORs of the Association between PET Risk Factors and PET incidence according to the reference standard data versus data obtained through the Ottawa hospital discharge database

Predictor Variables	Adjusted OR (PET reference standard)	p-value	Adjusted OR (PET hospital discharge database)	p-value
Ethnicity				
Black	2.43 (0.94-6.25)	$p=0.06$	1.60 (0.34-7.47)	$p=0.55$
Asian	1.68 (0.77 -3.63)	$p=0.18$	0.99 (0.23-4.26)	$p=0.99$
Caucasian (reference)	1.00		1.00	
Other	0.98 (0.53-1.81)	$p=0.95$	0.80 (0.30-2.12)	$p=0.66$
Chronic hypertension				
Yes	11.06 (5.37-22.78)	$p <.0001$	10.51 (3.97-27.82)	$p<.0001$
No (reference)	1.00		1.00	
Diabetes 1				
Yes	7.69 (3.05-19.34)	$p <.0001$	4.38 (0.97-19.82)	$p=0.06$
No (reference)	1.00		1.00	
Hx of PET				
Yes	6.57 (2.98-14.46)	$p<.0001$	15.12 (4.04-56.50)	$p<.0001$
No (reference)	1.00		1.00	
Parity				
Nullipara	2.73 (1.66-4.50)	$p<.0001$	10.32 (3.38-31.45)	$p<.0001$
Multipara (reference)	1.00		1.00	
Twin pregnancy				

Yes	6.10 (2.63-14.11)	p<.0001	3.12 (0.69-13.99)	p=0.14
No (reference)	1.00		1.00	
Smoking				
Yes	1.37 (0.60-3.09)	p=0.45	1.53 (0.45-5.17)	p=0.68
No (including quitting) (reference)	1.00		1.00	

In view of the directionality of misclassification, we found that even though certain parameters were underestimated in the hospital discharge database in comparison with those derived from the reference standard data, as expected for non-differential misclassification, there were also overestimated parameters in the hospital discharge database. One explanation for this phenomenon is that because the sensitivity of PET ascertainment is less than 50% and these errors occur non-differentially, the direction of effect can be reversed from the true situation. Another explanation is that there were other misclassified variables in the study. [56] Jurek et al. reported in their study that the belief that non-differential exposure misclassification always produces an underestimate of the true value is incorrect. One reason for this belief is a failure to understand that bias is not the only source of error in an estimate. Bias is not the ratio of the observed estimate from one study to the true value, because the observed estimate also incorporates random errors. Yet another reason for the incorrect belief is the failure to recognize that non-differentiality is insufficient to guarantee that the bias is towards the null; other conditions must be satisfied, especially independence of errors. Finally, even when bias due to exposure misclassification is towards the null, other biases (i.e., confounding, and mismeasurement of covariates), can cause the total bias to be away from the null. The combined effect from all biases must be considered when interpreting study results. [105]

6.6 Discussion

The idea of using administrative data for epidemiologic research is very attractive because of the easy access to thousands of records and the low cost; however, the validity of such data must be taken into consideration. The present study has demonstrated that the data from the hospital discharge database coded via the ICD-10-CA underestimated the incidence of PET. If the PET incidence obtained from such database is used in the predictive analysis, it can bias the results of the study. The study has shown further that if the studied factor is a strong predictor of PET, the extent of the association between the predictor and PET, would potentially depend on the source of PET ascertainment, though not significantly. On the other hand, if the studied factor is a new predictor which has not been well studied, then the extent of the association can be affected by the source of PET ascertainment.

We also found that certain parameters estimates were underestimated in the hospital discharge database in comparison with the parameters estimates derived from the reference standard data, as expected for non-differential misclassification; however, there were also parameter estimates that were overestimated in the hospital discharge database. To our knowledge, there has been only one study conducted evaluating the influence of misclassification on the strength of association. Klemmensen et al. conducted a validation study of the Danish National Hospital Registry using chart review as a gold standard. [76] They also compared the odds ratio for PET risk factors based on the gold standard with other sources of PET ascertainment and found that the odds ratios of the major risk factors showed the same tendency regardless of the source of diagnosis ascertainment. However, they also found that the odds ratio for one variable with controversial effect on

PET (smoking), according to the literature, depended on the database studied. When the diagnosis was obtained from the gold standard data, it tended to exhibit an increased risk of PET however; the estimates obtained from the administrative data tended to exhibit reduced risk of PET.

The consequences of PET misclassification depend on the association studied, the aim of the research, and other conditions. First of all, if the incidence of this condition is used to calculate the amount of money funded to a hospital for its research, the research of this serious condition of pregnancy might be seriously underfunded. Second, if the administrative data used to investigate the association of PET with health events later in life (for both, mother and offspring), the importance of accurate ascertainment of the disease for detecting the true associations is of utmost importance. Finally, the incidence of the data from administrative databases may be used to compare incidence rates among different hospitals, or even countries. Knowing that misclassification of the condition can be quite significant; the interpretation of the results should be done cautiously. All these factors should be taken into consideration before designing a research study based on the administrative data.

It should be noted that our study has certain limitations. It was done in only one hospital in Ontario. As a result, it is difficult to predict whether administrative data from a particular hospital, will underestimate or overestimate a given condition, unless a validation study is conducted. In addition, we used the adjudicated data from the OaK study as the “reference standard” and the PET definition used in the study was more inclusive than restrictive; consequently, these factors might suggest some caution in the interpretation of results.

6.7 Conclusion

The source of PET ascertainment should be taken into consideration when planning the research project; the extent of the association between the PET predictor and PET incidence would potentially be dependent on the source of PET ascertainment.

Chapter 7 General discussion

The purpose of this thesis was to uncover problems of the data derived from administrative databases based on the example of the Ottawa hospital discharge database and to propose ways to improve the quality of such data.

Administrative databases are used extensively because of their easy accessibility and the low cost. Accumulating evidence shows, however, that the accuracy of the data obtained through these databases is questionable; it does not reflect the true incidence of the disease studied.

The ICD is designed to promote international comparability in the collection, processing, classification, and presentation of mortality and morbidity statistics. The ICD-10 system has been implemented to allow better analysis of disease patterns and treatment outcomes in order to improve medical care. Coding accuracy is affected by the diagnosis definitions and the code interpretation. Therefore, it is relevant for PET - a condition with unclear etiology and pathogenesis, in particular. The ICD does not provide definitions for PET, although, there are a few definitions of PET used worldwide.

Our study conducted in one hospital in the province of Ontario has shown that PET incidence was significantly underreported in the Ottawa hospital discharge database in comparison with the reference standard which, in our study, was the adjudicated data

from the OaK Birth Cohort study. The adjudication process was validated through the calculation of the inter-observer agreement which was 0.82, or excellent agreement. [62]

We explored the potential causes of misclassification and developed a classification of errors. We found that the majority of the misclassified cases belonged to the category “PET pregnancies coded with incorrect ICD-10-CA code” (78.88%). The most significant finding was the diversity of diagnoses recorded in the medical charts which could be attributed to the confusion between two diagnoses: PIH and PET which were often used interchangeably. Even though, it has been recommended that the term PIH to be abandoned, as its meaning in clinical practice is unclear (SOGC recommendations) [2], we found that it is still widely used by clinicians. Additionally we explored the potential causes of misclassification. The most significant finding was the diversity of diagnoses recorded in the medical charts. We think that it could have contributed to PET misclassification due to the fact that two diagnoses, PIH and PET, are often used interchangeably.

We examined the extent to which the estimate of association between PET and some of the risk factors expressed in ORs has changed depending on whether PET incidence was based on the reference standard (adjudicated data) or the hospital discharge database (and ICD-10-CA). The results have demonstrated that strong predictors of PET, such as parity, history of PET, and pre-existing hypertension did not depend on the method of ascertaining PET. They all have shown strong association with PET regardless whether diagnosis was obtained through the reference standard data or through the hospital discharge database via ICD-10-CA code. We found that certain parameter estimates were underestimated in the hospital discharge database in comparison with the parameter

estimates derived from the reference standard data, as expected for non-differential misclassification, but there were also parameter estimates that were overestimated in the hospital discharge database.

The study focused on the accuracy of the hospital discharge database and ICD-10-CA for reporting of PET. The results of the study highlighted the low accuracy of the hospital discharge database for PET reporting. The degree of misclassification may make it difficult to detect the association that actually exists. By increasing awareness of potential sources of errors, users can better evaluate the applicability and limitations of the results derived from administrative databases.

We suggest that coding accuracy of PET might be improved through better education of coders and physicians. Currently CIHI is leading an improvement in the quality of health record coding and our results could be used to exemplify the importance of accurate coding. The physicians need to be educated about the broadening use of ICD codes (i.e., their use in surveillance activities) and the importance of explicit documentation in the discharge summary. The findings from studies such as our study need to be disseminated to clinicians. It should help improve future standardization and documentation.

We also suggest the incorporation of a clear definition of PET in the ICD-10-CA coding procedure manual. Clarification of the difference between two terms used interchangeably (PIH and PET) should also be provided.

We also propose that researchers using administrative databases in PET research should be aware of the limitations of such databases and should use additional sources of diagnosis ascertainment to improve the accuracy of diagnoses. Combination of different methods such as hospital discharge database and chart review might improve the

accuracy of PET diagnosis. Another way to improve the PET reporting is through creation of a disease (PET) specific database. A validation study to assess the accuracy of the data in the administrative database should also be considered before using these data in research, especially the one that has potential policy importance.

The study has limitations. First, retrospective design relied on data that had already been gathered. Certain variables included in the analysis were collected as self-reports. Even though an effort was made to verify self-reported information (e.g., smoking during pregnancy) against the information in the participant's medical records, there is no guarantee that all information in CRF is one hundred percent accurate. The chart review was also done retrospectively; therefore, we have no knowledge whether all information (clinical, discharge summaries and lab data) available to us was available to medical coders who worked in real time format.

Another limitation is lack of knowledge whether proper procedure was used for BP measurement. The proper procedure is a vital part of the hypertension diagnosis as per SOGC Guideline, but it was not documented in the medical charts.

Our study was done in only one hospital in the province of Ontario and, therefore, may not be representative of the coding process in other hospitals of the province or in Canada, in general. However, the incidence of PET obtained from the Ottawa hospital discharge database was similar to the incidence of PET according to the Niday Perinatal Database (Ontario Provincial database) (1.47% and 1.3% respectively) [87] and to the incidence of PET in one of the hospitals in the province of Alberta according to hospital discharge database (1.47% and 1.4% respectively) [80]. We believe that coding process

in other hospitals in Canada is similar to the coding process in the Ottawa Hospital, but more research needs to be done in this direction.

Finally, we used the adjudicated data from the OaK study as the “reference standard”, but PET definition used in adjudication was more inclusive than restrictive, and this may suggest that some caution be exercised in the interpretation of results. There are still debates among PET scholars about PET definition. According to some scholars, the definition should be simple, clinically useful, and allow predicting negative pregnancy outcomes. The problem is that clinicians are looking for a PET definition that would allow predicting poor obstetric outcomes, but epidemiology scholars are looking for a definition that would allow using the PET definition to determine PET trends and compare them among different hospitals in the same country or among different countries.

Chapter 8 Significance of the project

To our knowledge, there have only been a few previous studies conducted to examine the accuracy of administrative database via ICD-10-CA for diagnosis of PET, a major disease and one of the leading causes of maternal and perinatal morbidity and mortality.

Most of the research on the accuracy of hospital discharge databases in reporting obstetrical complications has been focused on hypertensive disorders in general not on PET, in particular. Also most of the research has been carried out in countries other than Canada.

To our knowledge this is the first Canadian hospital discharge database validation study conducted in the province of Ontario. Another Canadian study that validated hospital discharge database diagnosis through the Nova Scotia Atlee Perinatal Database showed similar results: low sensitivity for PET (although higher than in our study) with similar specificity (99.5%). [80]

Adjudication of outcomes conducted by a group of experts in the field, which is widely used in randomized controlled trials, is not common in observational research. This can be one of the reasons why the incidence of PET reported differs from study to study. The other reason, as reported by many scholars, [14, 16] could be the difference in PET definitions used among studies. Using adjudicated data as a “reference standard” to establish the accuracy of hospital discharge databases in reporting of PET is, to some extent, a validation study because it involves chart review. To increase the validity of the “reference standard” data, the inter-rater and agreement statistics was calculated.

To our knowledge, this is first study that has tried to explore the effect of basic patient identifiers on PET outcome misclassification explored the potential causes of

misclassification. The most significant finding was the diversity of diagnoses recorded in the medical charts which could be attributed to the confusion between two diagnoses: GH and PET which were often used interchangeably. We also developed a classification of errors and found that majority of the cases belong to the category “PET pregnancies coded with incorrect ICD-10-CA code” (78.88%).

The results of the study could be helpful for physicians and policy makers not only to obtain accurate data on incidence of preeclampsia but also to take into account quantified uncertainty of the results which can be used for further research, resource allocation and proposals to improve outcomes for women and their babies.

References

1. De Swiet, Michael. Medical Disorders in Obstetrical Practice. 4th Edition, 2002
2. Magee L. A., Helewa M., Moutquin J.-M., von Dadelszen P. "Diagnosis, Evaluation, and Management of the Hypertensive Disorders of Pregnancy." *Journal of Obstetrics and Gynaecology Canada*. (2008); 30; 3.
3. Dekker G, Sibai B. "Primary, secondary, and tertiary prevention of pre-eclampsia." *Lancet*. (2001); 357; 209-215.
4. Dolea C., AbouZahr C. "Global burden of hypertensive disorders of pregnancy in the year 2000." *WHO report 2000*.
5. Baumwell S, Karumanchi SA. "Pre-Eclampsia: Clinical Manifestations and Molecular Mechanisms." *Nephron Clinical Practice*. (2007); 106; c72-c81.
6. Hawfield A., Freedman B. I. "Pre-eclampsia: the pivotal role of the placenta in its pathophysiology and markers for early detection." *Therapeutic Advances in Cardiovascular Disease*. (2009); 3(1); 65–73.
7. Kaaja R. "Predictors and risk factors of pre-eclampsia." *Minerva Ginecologica*. (2008); 60(5); 421-9.
8. World Health Organization, "Global Program to Conquer Preeclampsia/Eclampsia." 2002.
9. Duley L. "The Global Impact of Pre-eclampsia and Eclampsia." *Seminars in Perinatology*. (2009); 33; 3; 130-137.
10. Nelson Textbook of Pediatrics. Behrman; 2004.
11. Garovic V.D., Hayman S.R. "Hypertension in pregnancy: an emerging risk factor for cardiovascular disease." *Nature*. (2007); 3; 613–622.

12. Ray J.G., Vermeulen M.J., Schull M.J., Redelmeier D.A. "Cardiovascular health after maternal placental syndromes (CHAMPS): population-based retrospective cohort study." *Lancet*. (2005); 366; 1797–1803.
13. Harlow F. H., Brown M. A. "The diversity of diagnoses of preeclampsia. Hypertension in Pregnancy." *Hypertension and Pregnancy*. (2001); 20(1); 57-67.
14. American College of Obstetricians and Gynecologists. "ACOG practice bulletin. Diagnosis and management of preeclampsia and eclampsia. Number 33, January 2002." *Obstetrics & Gynecology*. (2002); 99(1); 159-67.
15. National Institutes of Health; National Heart, Lung, and Blood Institute; National High Blood Pressure Education Program. Report of the Working Group on Research on Hypertension During Pregnancy. (2001); <http://www.nhlbi.nih.gov/>.
16. Brown M. A., de Swiet M. "Classification of hypertension in Pregnancy". *Baillieres Best Practice & Research Clinical Obstetrics & Gynaecology*. (1999); 13(1); 27-39.
17. Villar J, Say L, Shennan A, Lindheimer M, Duley L, Conde-Agudelo A, Merialdi M. "Methodological and technical issues related to the diagnosis, screening, prevention, and treatment of pre-eclampsia and eclampsia." *International Journal of Gynecology and Obstetrics*. (2004); 85 Suppl 1; S28-41.
18. Davey DA, MacGillivray I. "The classification and definition of the hypertensive disorders of pregnancy." *American Journal of Obstetrics & Gynecology*. (1988); 158(4): 892-8.
19. Australasian Society for the Study of Hypertension in Pregnancy. "Management of Hypertension in Pregnancy: Consensus Statement." *The Medical Journal of Australia*. (1993); 158; 700-702.

20. National High Blood Pressure Education Program. Working Group Report on High Blood Pressure in Pregnancy. *American Journal of Obstetrics & Gynecology* (1990); 163: 1689-1712.
21. Helewa M. E., Burrows R. F., Smith J., Williams K., Brain P., Rabkin S. W. "Report of the Canadian Hypertension Society Consensus Conference: 1. Definitions, evaluation and classification of hypertensive disorders in pregnancy." *Canadian Medical Association Journal*. (1997); 157(6): 715–725.
22. Yasmeeen S., Romano P.S., Schembri M.E., Keyzer J.M., Gilbert W.M. "Accuracy of obstetric diagnoses and procedures in hospital discharge data." *American Journal of Obstetrics and Gynecology*. (2006); 194, 992–1001.
23. Quan H., Parsons G. A., Ghali W. A. "Validity of Information on Comorbidity Derived From ICD-9-CCM Administrative Data." *Medical Care*. (2002); 40; 8; 675-685.
24. CIHI - Taking Information Further. (2005). <http://secure.cihi.ca/cihiweb>.
25. CIHI - Discharge Abstract Database (DAD)/ CMG/Plx Data Quality Re-abstraction Study. (2003).
26. McMahon L.F., Jr., Smits H.L. "Can Medicare prospective payment survive the ICD-9-CM disease classification system?" *Ann Intern Med*. (1986); 104(4); 562-566.
27. World Health Organization. "International classification of disease and related health problems, tenth revision (ICD-10)". Geneva, 1992.
28. Walker S., Wood M., Wilks J., Nicol J. "Comparing ICD-9-CM and ICD-10 classification systems in a primary health care setting: some initial observations." *Health Information Management Journal*. (1995); 25(3); 83-86.

29. No authors listed. "Comparison of ICD-9-CM and ICD-10-AM. Australian modifications; ICD-10 changes." *Health Information Management Journal*. (1997); 27(1); 23-27.
30. Canadian Institute for Health Information. International statistical classification of disease and related health problems, tenth revision, Canada (ICD-10-CA). Ottawa, 2003.
31. Cheng P., Gilchrist A., Robinson K. M., Paul L. "The risk and consequences of clinical miscoding due to inadequate medical documentation: a case study of the impact on health services funding." *Health Information Management Journal*. (2009); 38; 1; pp 35-46.
32. Marshall J., Adema D. "Reinventing radiology reimbursement." *Radiology Management*. (2005); 27(2); 36-44, 46.
33. Brown D. W., Croft J. B., Greenlund, K. J., Giles, W.H. "Deaths from chronic obstructive pulmonary disease-United States." *Morbidity and Mortality Weekly report*. (2008); 57(45); 1229-1232.
34. Schoenman, J., Sutton J. P., Elixhopuser A., Love D. "Understanding and enhancing the value of hospital discharge data." *Medical Care Research and Review*. (2008); 64(4); 449-468.
35. Muhajarine N., Mustard C., Roos L.L., Young T.K., Gelskey D.E. "Comparison of survey and physician claims data for detecting hypertension." *Journal of Clinical Epidemiology*. (1997); 50(6); 711-8.

36. Chiò A., Ciccone G., Calvo A., Vercellino M., Di Vito N., Ghiglione P., Mutani R.
“Validity of hospital morbidity records for amyotrophic lateral sclerosis. A population-based study.” *Journal of Clinical Epidemiology*. (2002); 55; 723-727.
37. Tairou F., De Wals P., Bastide A. “Validity of death and stillbirth certificates and hospital discharge summaries for the identification of neural tube defects in Quebec City.” *Chronic Disease in Canada*. (2006); 27; 3; 120-124.
38. Dodds L., Spencer A., Shea S., Fell D., Armson B. A., Allen A. C., Bryson S..
“Validity of autism diagnoses using administrative health data.” *Chronic Disease in Canada*. (2009); 29; 3; 102-107.
39. Quan H., Li B., Saunders L D., Parsons G. A, Nilsson C. I., Alibhai A., Ghali W. A.
“Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database.” *Health Services Research*. (2008); 43; 4; 1424-1441.
40. Humphries K.H., Rankin J.M., Carere R.G., Buller C.E., Kiely F.M., Spinelli J.J.
“Co-morbidity data in outcome research. Are clinical data derived from administrative database a reliable alternative to chart review?” *Journal of Clinical Epidemiology*. (2000); 53; 343-349.
41. Ellekjær H., Holmen J., Krüger Ø., Terent A. “Identification of Incident Stroke in Norway Hospital Discharge Data Compared With a Population-Based Stroke Register.” *Stroke*. (1999); 30; 56-60.
42. Lofthus C.M., Cappelen I., Osnes E.K., Falch J.A., Kristiansen I.S., Medhus A.W., Nordsletten L., Meyer H.E. “Local and national electronic databases in Norway

demonstrate a varying degree of validity.” *Journal of Clinical Epidemiology*. (2005); 58; 280–285.

43. Larsen T.B., Johnsen S.P., Møller C.I., Larsen H., Sørensen H.T. “A review of medical records and discharge summary data found moderate to high predictive values of discharge diagnoses of venous thromboembolism during pregnancy and postpartum.” *Journal of Clinical Epidemiology*. (2005); 58; 316–319.
44. Rawson, N.S.B., D’Arcy, C. “Validity and reliability: idealism and reality in the use of computerized health care databases for pharmacoepidemiological research.” *Post Marketing Surveillance*. (1991); 5; 31-55.
45. Stergachis, A.C. “Evaluating the quality of linked automated databases for use in pharmacoepidemiology.” *Pharmacoepidemiology: An Introduction*, 2nd edition. Harvey Whitney, Cincinnati, 1991; 222-234.
46. Rawson, N.S.B., D’Arcy, C. “Assessing the Validity of Diagnostic Information in Administrative Health Care Utilization Data: Experience in Saskatchewan.” *Pharmacoepidemiology and drug safety*. (1998); 7; 389-398.
47. De Coster C., Quan H., Finlayson A., Gao M., Halfon P., Humphries K. H., Johansen H., Lix L. M., Luthi J.-C., Ma J., Romano P. S., Roos L., Sundararajan V., Tu J. V., Webster G., Ghali W. A. “Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium.” *BMC Health Services Research* (2006); 6:77.
48. ACOG technical bulletin. Management of preeclampsia. No. 91—February 1986. Committee on Technical Bulletins of the American College of Obstetricians and Gynecologists. Washington, D.C., 1986.

49. Cunningham F.G. Williams Obstetrics 19th Edition, 1993.
50. Brown M. A., Buddle M. L. "What's in a name? Problems with the classification of hypertension in pregnancy." *Journal of Hypertension*. (1997); 15; 10; 1049-1054.
51. Helewa M.E., Burrows R.F., Smith J., Williams K., Brain P., Rabkin S.W. "Report of the Canadian Hypertension Society Consensus Conference: 1. Definitions, evaluation and classification of hypertensive disorders in pregnancy." *Canadian Medical Association Journal*. (1997); 15; 157(6); 715-25.
52. Chappell L., Poulton L., Halligan A., Shennan A.H. "Lack of consistency in research papers over the definition of pre-eclampsia." *British Journal of Obstetrics and Gynaecology*. (1999); 106; 983-985.
53. Higgins J.R., de Swiet M. "Blood-pressure measurement and classification in pregnancy." *Lancet*. (2001); 357; 131-135.
54. North R.A., Taylor R.S., Schellenberg J.C. "Evaluation of a definition of pre-eclampsia." *British Journal of Obstetrics and Gynaecology*. (1999); 1; 106, 767-773.
55. The Canadian Enhancement of ICD-10 (International Statistical Classification of Diseases and Related Health Problems, Tenth Revision). CIHI 2001.
56. Rothman K. J., Greenland S., Lash T. L. Modern Epidemiology, 3rd Edition, 2008.
57. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP). 2005.
<http://www.emea.europa.eu/pdfs/human/ewp/587203en.pdf>.
58. US Food and Drug Administration. Guidance for clinical trial sponsors on the establishment and operation of clinical trial data monitoring committees. 2001.

<http://www.fda.gov/CBER/gdlns/clindatmon.pdf>.

59. Mahaffey K.W., Harrington R.A., Akkerhuis M., Kleiman N.S., Berdan L.G., Crenshaw B.S., Tardiff B.E., Granger C.B., DeJong I., Bhapkar M., Widimsky P., Corbalon R., Lee K.L., Deckers J.W., Simoons M.L., Topol E.J., Califf R.M. “Systematic adjudication of myocardial infarction end-point in an international clinical trial.” *Current controlled trials in cardiovascular medicine*. (2001); 1; 4; 180-186.
60. Mahaffey K.W., Harrington R.A., Akkerhuis M., Kleiman N.S., Berdan L.G., Crenshaw B.S., Tardiff B.E., Granger C.B., DeJong I., Bhapkar M., Widimsky P., Corbalon R., Lee K.L., Deckers J.W., Simoons M.L., Topol E.J., Califf R.M. “Disagreement between central clinical event committee and site investigators assessment infarction end-points in an international clinical trial: review of the PURSUIT study.” *Current controlled trials in cardiovascular medicine*. (2001); 2; 187-194.
61. Pogue J., Walter S.D., Yusuf S. “Evaluating the benefit of event adjudication of cardiovascular outcomes in large simple RCTs.” *Clinical Trials*. (2009); 6; 239-251.
62. Heckbert S.R., Kooperberg C., Safford M.M., Psaty B.M., Hsia J., McTiernan A., Gaziano J.M., Frishman W.H., Curb J.D. “Comparison of Self-Report, Hospital Discharge Codes, and Adjudication of Cardiovascular Events in the Women’s Health Initiative.” *American Journal of Epidemiology*. (2004); 160; 1152–1158.
63. Cohen J. “A coefficient of agreement for nominal scales.” *Educational and Psychological Measurement*. (1960); 20; 27-46.

64. Kundel H. L., Polansky M. "Measurement of Observer Agreement." *Radiology*. (2003); 228; 303–308.
65. Fleiss J.L. *Statistical Methods for Rates and Proportions*, John Wiley and Sons, New York, 1981.
66. Landis J.R., Koch G.G. "The Measurement of Observer Agreement for Categorical Data." *Biometrics*. (1977); 33; 159-174.
67. Dechartres A., Boutron I., Roy C., Ravaud P. "Inadequate planning and reporting of adjudication committees in clinical trials. Recommendation proposal." *Journal of Clinical Epidemiology*. (2009); 62; 695-702.
68. Roberts C.L., Bell J.C., Ford J.B., Hadfield R.M., Algert C.S., Morris J.M. "The Accuracy of Reporting of the Hypertensive Disorders in Pregnancy in Population Health Data." *Hypertension in Pregnancy*. (2008); 27; 3; 285-297.
69. O'Malley, K. J., Cook K. F., Price M. D., Wildes K. R., Hurdle J. F., Ashton C. M. "Measuring Diagnoses: ICD Code Accuracy." *Health Services Research*. (2005); 40; 5; 1620-1639.
70. Ozalp S., Tanir H.M., Sener T., Yazan S., Keskin A.E. "Health risks for early (< or =19) and late (> or =35) childbearing." *Archives of Gynecology and Obstetrics*. (2003); 268; 172-174.
71. Luck J., Peabody J.W., Dresselhaus T.R., Lee M., Glassman P. "How Well Does Chart Abstraction Measure Quality? A Prospective Comparison of Standardized Patients with the Medical Record." *The American Journal of Medicine*. (2000); 108; 642-649.

72. Henderson T., Shephard J., Sundararajan V. "Quality of diagnosis and procedure coding in ICD-10 code administrative data." *Medical Care*. (2006); 44; 1011-1019.
73. Thornton C., Makris A., Ogle R., Hennessy A. "Generic obstetrics database systems are unreliable for reporting the hypertensive disorders of pregnancy." *Australian and New Zealand Journal of Obstetrics and Gynaecology*. (2004); 44; 505-509.
74. Roberts C. L., Algert C. S., Ford J. B. "Methods for dealing with discrepant records in linked population health datasets: a cross-sectional study." *BMS Health Services Research*. (2007); 7:12.
75. Geller S.E., Ahmed S., Brown M.L., Cox S.M., Rosenberg D., Kilpatrick S.J. "International Classification of Diseases-9th Revision coding for preeclampsia: How accurate is it?" *American Journal of Obstetrics and Gynecology*. (2004); 190; 1629-34.
76. Klemmensen Å. K., Olsen S F., Østerdal M. L., Tabor A. "Validity of Preeclampsia-related Diagnoses Recorded in a National Hospital Registry and in a Postpartum Interview of the Women." *American Journal of Epidemiology*. (2007); 166; 117-124
77. Callaghan W. M. "Invited Commentary: Identifying Women with Hypertension during Pregnancy-is High Specificity Sufficient?" *American Journal of Epidemiology*. (2007); 166; 2; 125-127.
78. Klemmensen Å. K., Olsen S.F., Wengel C.M., Tabor A. "Diagnostic criteria and reporting procedures for pre-eclampsia: A national survey among obstetrical departments in Denmark." *European Journal of Obstetrics & Gynecology and Reproductive Biology*. (2005); 123; 41-45

79. Joseph K.S., Fahey J. "Validation of perinatal data in the Discharge abstract Database of the Canadian Institute for Health Information." *Chronic Disease in Canada*. (2009); 29; 3, 96-100.
80. Walker R. L., Hemmelgarn B., Quan H. "Incidence of gestational hypertension in the Calgary Health Region from 1995 to 2004." *The Canadian Journal of Cardiology*. (2009); 25; 8; e284-e287.
81. Roberts C. L., Ford J. B., Algert C. S., Antonsen S., Chalmers J., Cnattingius S, Gokhale M., Kotelchuck M., Melve K. K., Langridge A., Morris C., Morris J. M., Nassar N., Norman J. E., Norrie J., Sørensen H. T., Walker R., Weir C. J. "Population-based trends in pregnancy hypertension and pre-eclampsia: an international comparative study." *BMJ Open*. (2011). doi:10.1136/bmjopen-2011-000101.
82. Flahault A., Cadilhaca M., Thomas G. "Sample size calculation should be performed for design accuracy in diagnostic test studies." *Journal of Clinical Epidemiology*. (2005); 58; 8, 859-862.
83. Brubaker, L., Bradley, C. S., Handa, V. L., Richter, H. E., Visco, A., Brown, M. B., Weber, A.M. "Anal sphincter laceration at vaginal delivery: is this event coded accurately?" *Obstetrics and Gynecology*. (2007); 109(5); 1141-1145.
84. O'Malley, K. J., Cook, K. F., Price, M. D., Wildes, K. R., Hurdle, J. F., Ashton, C. M. "Measuring diagnoses: ICD code accuracy." *Health Services Research*. (2005); 40(5 Pt 2); 1620-1639.
85. Surjan G. "Questions on validity of International Classification of Disease-coded diagnoses." *International Journal of Medical Informatics*. (1999); 54; 77-95.

86. Canadian Institutes of Health Research, Natural Sciences and Engineering Research Council of Canada, Social Sciences and Humanities Research Council of Canada, Tri-council Policy Statement: Ethical Conduct for Research Involving Humans, 1998 (with 2000, 2002, 2005 amendments), 93 p.
87. Ontario Perinatal Surveillance System Committee. The Ontario Perinatal Surveillance System Report 2008. Ottawa, ON: Ontario Perinatal Surveillance System; 2008.
88. Saftlas A.F., Olson D.R., Franks A.L., Atrash H.K., Pokras R. "Epidemiology of preeclampsia and eclampsia in the United states, 1979-1986." *American Journal of Obstetrics and Gynecology*. (1990); 163: 460-5.
89. Coonrod D.V., Hickok D.E., Zhu K., Easterling T.R., Daling J.R. "Risk factors for preeclampsia in twin pregnancies: a population-based cohort study." *Obstetrics and Gynecology*. (1995); 85; 645-50.
90. Lawoyin T.O., Ani F. "Epidemiologic aspects of pre-eclampsia in Saudi Arabia." *The East African Medical Journal*. (1996); 73; 404-6.
91. Lee C.-J., Hsieh T.-T., Chiu T.-H., Chen K.-C., Lo L.-M., Hung T.-H. "Risk factors for pre-eclampsia in an Asian population." *International Journal of Gynecology & Obstetrics*. (2000); 70; 327-3.
92. Makkonen N., Heinonen S., Kirkinen P. "Obstetric prognosis in second pregnancy after preeclampsia in first pregnancy." *Hypertension and Pregnancy*. (2000); 19; 173-81.
93. Dukler D., Porath A., Bashiri A., Erez O., Mazor M. "Remote prognosis of primiparous women with preeclampsia." *European Journal of Obstetrics & Gynecology and Reproductive Biology*. (2001); 96; 69-74.

94. Odegard R.A., Vatten L.J., Nilsen S.T., Salvesen K.A., Austgulen R. "Risk factors and clinical manifestations of pre-eclampsia." *British Journal of Obstetrics & Gynaecology*. (2000); 107; 1410-6.
95. Santema J.G., Koppelaar I., Wallenburg H.C. "Hypertensive disorders in twin pregnancy." *European Journal of Obstetrics & Gynecology and Reproductive Biology*. (1995); 58; 9-13.
96. Garner P.R., D'Alton M.E., Dudley D.K., Huard P., Hardie M. "Preeclampsia in diabetic pregnancies." *American Journal of Obstetrics & Gynecology*. (1990); 163; 505-8.
97. Ros H.S., Cnattingius S., Lipworth L. "Comparison of risk factors for preeclampsia and gestational hypertension in a population-based cohort study." *The American Journal of Epidemiology*. (1998); 147; 11(suppl); 1062-70.
98. Rey E., Couturier A. "The prognosis of pregnancy in women with chronic hypertension." *American Journal of Obstetrics & Gynecology*. (1994); 171; 410-6.
99. Sibai B.M., Ewell M., Levine R.J., Klebanoff M.A., Esterlitz J., Catalano P.M., Goldenberg R.L., Joffe G. "Risk factors associated with preeclampsia in healthy nulliparous women. The Calcium for Preeclampsia Prevention (CPEP) Study Group." *American Journal of Obstetrics & Gynecology*. (1997); 177; 1003-10.
100. Leeners, B., Neumaier-Wagner, P., Kuse, S. & Rath, W. "Smoking and the risk of developing hypertensive diseases in pregnancy: what is the effect on HELLP syndrome?" *Acta Obstetrica et Gynecologica Scandinavica*. (2006); 85; 1217-1224.

101. Conde-Agudelo A., Althabe F., Belizán J.M., Kafury-Goeta A.C. "Cigarette smoking during pregnancy and risk of preeclampsia: A systematic review." *American Journal of Obstetrics & Gynecology*. (1999); 181; 4; 1026-1035.
102. England L., Zhang J. "Smoking and risk of preeclampsia: a systematic review." *Frontiers in Bioscience*. (2007); 12; 2471-2483.
103. England L.J., Levine R.J., Qian C., Morris C.D., Sibai B.M., Catalano P.M., Curet L.B., Klebanoff M.A. "Smoking before pregnancy and risk of gestational hypertension and preeclampsia." *American Journal of Obstetrics & Gynecology*. (2002); 186; 5; 1035-1040.
104. Tanaka M., Jaamaa G., Kaiser M., Hills E., Soim A., Zhu M., Shcherbatykh I.Y., Samelson R., Bell E., Zdeb M., McNutt L.A. "Racial disparity in hypertensive disorders of pregnancy in New York State: a 10-year longitudinal population-based study." *The American Journal of Public Health*. (2007); 97; 163-170.
105. Jurek M., Greenland S., Maldonado G., Church T.R. "Proper interpretation of non-differential misclassification effects: expectations vs observations." *International Journal of Epidemiology*. (2005); 34; 680–687.
106. Belsley D.A., Kuh E. Welsch R.E. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley & Sons: New York 1980.
107. Wassertheil-Smoller S. *Biostatistics and Epidemiology: A Primer for Health and Biomedical Professionals*. Third Edition 2004.
108. Hosmer D.W., Lemeshow S. *Applied Logistic Regression*. John Wiley & Sons: New York 1989.

109. Preventing and Managing the Global Epidemic of Obesity. Report of the World Health Organization Consultation of Obesity. WHO, Geneva, June 1997.

Appendices

Appendix 1 Classification of hypertensive diseases of pregnancy according to ICD-10-CA

Oedema, proteinuria and hypertensive disorders in pregnancy, childbirth and the puerperium (O10-O16)

Note: In order to satisfy the needs of the Canadian grouping methodology, appropriate codes in the range O10.0 to O99.8 will have a supplementary subclassification placed at the end of the code to identify the following

- 1 **Delivered, with or without mention of antepartum condition**

- Antepartum condition with delivery - Delivery NOS - Intrapartum obstetric condition - Pregnancy delivered	}	with or without mention of antepartum or intrapartum complication during the current episode of care
--	---	--
- 2 **Delivered with mention of postpartum**

- Delivery - Pregnancy, delivered	}	with mention of postpartum or puerperal complication during the current episode of care
--	---	---
- 3 **Antepartum condition or complication**
Antepartum obstetric condition, not delivered during the current episode of care
- 4 **Postpartum condition or complication**
Postpartum or puerperal obstetric condition or complication following delivery that occurred
 - During previous episode of care
 - Outside hospital with subsequent admission for observation or care
- 9 **Unspecified as to episode of care**

O10 Pre-existing hypertension complicating pregnancy, childbirth and the puerperium

Includes: the listed conditions with pre-existing proteinuria

Excludes: that with increased or superimposed proteinuria (O11)

O10 Pre-existing hypertension complicating pregnancy, childbirth and the puerperium	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum condition	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O10.0 Pre-existing essential hypertension complicating pregnancy, childbirth and the puerperium Includes • Any condition in I10 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.001*	O10.002*	O10.003*	O10.004*	O10.009*

O10 Pre-existing hypertension complicating pregnancy, childbirth and the puerperium	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum condition	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O10.1 Pre-existing hypertensive heart disease complicating pregnancy, childbirth and the puerperium Includes • Any condition in I11 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.101✦	O10.102✦	O10.103✦	O10.104✦	O10.109✦
O10.2 Pre-existing hypertensive renal disease complicating pregnancy, childbirth and the puerperium Includes • Any condition in I12 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.201✦	O10.202✦	O10.203✦	O10.204✦	O10.209✦
O10.3 Pre-existing hypertensive heart and renal disease complicating pregnancy, childbirth and the puerperium Includes • Any condition in I13 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.301✦	O10.302✦	O10.303✦	O10.304✦	O10.309✦
O10.4 Pre-existing secondary hypertension complicating pregnancy, childbirth and the puerperium Includes • Any condition in I15.- specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.401✦	O10.402✦	O10.403✦	O10.404✦	O10.409✦
O10.9 Unspecified pre-existing hypertension complicating pregnancy, childbirth and the puerperium	O10.901✦	O10.902✦	O10.903✦	O10.904✦	O10.909✦

O11

Pre-existing hypertensive disorder with superimposed proteinuria

Includes: Conditions in O10.- complicated by increased proteinuria
Superimposed pre-eclampsia

- O11.001✦ Pre-existing hypertensive disorder with superimposed proteinuria, delivered, with or without mention of antepartum condition
- O11.002✦ Pre-existing hypertensive disorder with superimposed proteinuria, delivered, with mention of postpartum complication
- O11.003✦ Pre-existing hypertensive disorder with superimposed proteinuria, antepartum condition or complication

- O11.004 ✦ Pre-existing hypertensive disorder with superimposed proteinuria, postpartum condition or complication
- O11.009 ✦ Pre-existing hypertensive disorder with superimposed proteinuria, unspecified as to episode of care, or not applicable

O12 Gestational [pregnancy-induced] oedema and proteinuria without hypertension

O12 Gestational [pregnancy-induced] oedema and proteinuria without hypertension	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum complication	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O12.0 Gestational oedema	O12.001 ✦	O12.002 ✦	O12.003 ✦	O12.004 ✦	O12.009 ✦
O12.1 Gestational proteinuria	O12.101 ✦	O12.102 ✦	O12.103 ✦	O12.104 ✦	O12.109 ✦
O12.2 Gestational oedema with proteinuria	O12.201 ✦	O12.202 ✦	O12.203 ✦	O12.204 ✦	O12.209 ✦

O13 Gestational [pregnancy-induced] hypertension without significant proteinuria

Includes: Gestational hypertension NOS
Mild pre-eclampsia
Transient hypertension of pregnancy

- O13.001 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, delivered, with or without mention of antepartum condition
- O13.002 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, delivered, with mention of postpartum complication
- O13.003 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, antepartum condition or complication
- O13.004 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, postpartum condition or complication
- O13.009 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, unspecified as to episode of care, or not applicable

O14 Gestational [pregnancy-induced] hypertension with significant proteinuria

Includes: HELLP (syndrome) (hemolysis/elevated liver enzymes/low platelets)
Excludes: superimposed pre-eclampsia (O11)

- O14.001 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, delivered, with or without mention of antepartum condition
- O14.002 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, delivered, with mention of postpartum complication
- O14.003 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, antepartum condition or complication
- O14.004 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, postpartum condition or complication
- O14.009 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, unspecified as to episode of care, or not applicable

O15**Eclampsia**

Includes: convulsions following conditions in O10-O14 and O16
 eclampsia with pregnancy-induced or pre-existing hypertension
 Use additional code to identify any associated status epilepticus (G41.0)

O15 Eclampsia	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum condition	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O15.0 Eclampsia in pregnancy	O15.001*	---	O15.003*	---	---
O15.1 Eclampsia in labour	O15.101*	---	O15.103*	---	---
O15.2 Eclampsia in the puerperium	---	O15.202*	---	O15.204*	---
O15.9 Eclampsia, unspecified as to time period <i>Includes</i> • Eclampsia NOS	---	---	---	---	O15.909*

O16**Unspecified maternal hypertension**

- O16.001* Unspecified maternal hypertension, delivered, with or without mention of antepartum condition
- O16.002* Unspecified maternal hypertension, delivered, with mention of postpartum complication
- O16.003* Unspecified maternal hypertension, antepartum condition or complication
- O16.004* Unspecified maternal hypertension, postpartum condition or complication
- O16.009* Unspecified maternal hypertension, unspecified as to episode of care, or not applicable

Oedema, proteinuria and hypertensive disorders in pregnancy, childbirth and the puerperium (O10-O16)

Note: In order to satisfy the needs of the Canadian grouping methodology, appropriate codes in the range O10.0 to O99.8 will have a supplementary subclassification placed at the end of the code to identify the following

- 1 **Delivered, with or without mention of antepartum condition**

- Antepartum condition with delivery - Delivery NOS - Intrapartum obstetric condition - Pregnancy delivered	}	with or without mention of antepartum or intrapartum complication during the current episode of care
--	---	--
- 2 **Delivered with mention of postpartum**

- Delivery - Pregnancy, delivered	}	with mention of postpartum or puerperal complication during the current episode of care
--	---	---
- 3 **Antepartum condition or complication**
Antepartum obstetric condition, not delivered during the current episode of care
- 4 **Postpartum condition or complication**
Postpartum or puerperal obstetric condition or complication following delivery that occurred
 - During previous episode of care
 - Outside hospital with subsequent admission for observation or care
- 9 **Unspecified as to episode of care**

O10 Pre-existing hypertension complicating pregnancy, childbirth and the puerperium

Includes: the listed conditions with pre-existing proteinuria

Excludes: that with increased or superimposed proteinuria (O11)

O10 Pre-existing hypertension complicating pregnancy, childbirth and the puerperium	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum condition	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O10.0 Pre-existing essential hypertension complicating pregnancy, childbirth and the puerperium Includes • Any condition in I10 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.001*	O10.002*	O10.003*	O10.004*	O10.009*

O10 Pre-existing hypertension complicating pregnancy, childbirth and the puerperium	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum condition	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O10.1 Pre-existing hypertensive heart disease complicating pregnancy, childbirth and the puerperium Includes • Any condition in I11 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.101✳	O10.102✳	O10.103✳	O10.104✳	O10.109✳
O10.2 Pre-existing hypertensive renal disease complicating pregnancy, childbirth and the puerperium Includes • Any condition in I12 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.201✳	O10.202✳	O10.203✳	O10.204✳	O10.209✳
O10.3 Pre-existing hypertensive heart and renal disease complicating pregnancy, childbirth and the puerperium Includes • Any condition in I13 specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.301✳	O10.302✳	O10.303✳	O10.304✳	O10.309✳
O10.4 Pre-existing secondary hypertension complicating pregnancy, childbirth and the puerperium Includes • Any condition in I15.- specified as a reason for obstetric care during pregnancy, childbirth or the puerperium	O10.401✳	O10.402✳	O10.403✳	O10.404✳	O10.409✳
O10.9 Unspecified pre-existing hypertension complicating pregnancy, childbirth and the puerperium	O10.901✳	O10.902✳	O10.903✳	O10.904✳	O10.909✳

O11

Pre-existing hypertensive disorder with superimposed proteinuria

Includes: Conditions in O10.- complicated by increased proteinuria
Superimposed pre-eclampsia

- O11.001✳ Pre-existing hypertensive disorder with superimposed proteinuria, delivered, with or without mention of antepartum condition
- O11.002✳ Pre-existing hypertensive disorder with superimposed proteinuria, delivered, with mention of postpartum complication
- O11.003✳ Pre-existing hypertensive disorder with superimposed proteinuria, antepartum condition or complication

- O11.004 ✦ Pre-existing hypertensive disorder with superimposed proteinuria, postpartum condition or complication
- O11.009 ✦ Pre-existing hypertensive disorder with superimposed proteinuria, unspecified as to episode of care, or not applicable

O12 Gestational [pregnancy-induced] oedema and proteinuria without hypertension

O12 Gestational [pregnancy-induced] oedema and proteinuria without hypertension	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum complication	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O12.0 Gestational oedema	O12.001 ✦	O12.002 ✦	O12.003 ✦	O12.004 ✦	O12.009 ✦
O12.1 Gestational proteinuria	O12.101 ✦	O12.102 ✦	O12.103 ✦	O12.104 ✦	O12.109 ✦
O12.2 Gestational oedema with proteinuria	O12.201 ✦	O12.202 ✦	O12.203 ✦	O12.204 ✦	O12.209 ✦

O13 Gestational [pregnancy-induced] hypertension without significant proteinuria

Includes: Gestational hypertension NOS
Mild pre-eclampsia
Transient hypertension of pregnancy

- O13.001 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, delivered, with or without mention of antepartum condition
- O13.002 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, delivered, with mention of postpartum complication
- O13.003 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, antepartum condition or complication
- O13.004 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, postpartum condition or complication
- O13.009 ✦ Gestational [pregnancy-induced] hypertension without significant proteinuria, unspecified as to episode of care, or not applicable

O14 Gestational [pregnancy-induced] hypertension with significant proteinuria

Includes: HELLP (syndrome) (hemolysis/elevated liver enzymes/low platelets)
Excludes: superimposed pre-eclampsia (O11)

- O14.001 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, delivered, with or without mention of antepartum condition
- O14.002 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, delivered, with mention of postpartum complication
- O14.003 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, antepartum condition or complication
- O14.004 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, postpartum condition or complication
- O14.009 ✦ Gestational [pregnancy-induced] hypertension with significant proteinuria, unspecified as to episode of care, or not applicable

O15**Eclampsia**

Includes: convulsions following conditions in O10-O14 and O16
 eclampsia with pregnancy-induced or pre-existing hypertension
 Use additional code to identify any associated status epilepticus (G41.0)

O15 Eclampsia	Delivered, with or without mention of antepartum condition	Delivered, with mention of postpartum condition	Antepartum condition or complication	Postpartum condition or complication	Unspecified as to episode of care, or not applicable
O15.0 Eclampsia in pregnancy	O15.001*	---	O15.003*	---	---
O15.1 Eclampsia in labour	O15.101*	---	O15.103*	---	---
O15.2 Eclampsia in the puerperium	---	O15.202*	---	O15.204*	---
O15.9 Eclampsia, unspecified as to time period <i>Includes</i> * Eclampsia NOS	---	---	---	---	O15.909*

O16**Unspecified maternal hypertension**

- O16.001* Unspecified maternal hypertension, delivered, with or without mention of antepartum condition
- O16.002* Unspecified maternal hypertension, delivered, with mention of postpartum complication
- O16.003* Unspecified maternal hypertension, antepartum condition or complication
- O16.004* Unspecified maternal hypertension, postpartum condition or complication
- O16.009* Unspecified maternal hypertension, unspecified as to episode of care, or not applicable

Appendix 2 Adjudication Extra Form



43932



Recruitment Number

Page 1 of 1

HYPERTENSION IN PREGNANCY

1. Chronic Hypertension (pre-existing hypertension) ☐ Yes ☐ No

◆ Present and observable prior to pregnancy or is diagnosed before the 20th week of gestation

◆ BP greater than 140 mmHg systolic &/OR 90 mmHg diastolic

◆ Persists > 42 days postpartum

2. Hypertension of Pregnancy ☐ Yes ☐ No IF 'YES' Please complete this page

◆ BP greater than 140 &/OR 90 mmHg which develops after 20th week of gestation

Diastolic $\geq$ 90 mmHg or an increase of 15 mmHg *from the participant's baseline

Systolic $\geq$ 140 mmHg or an increase of 30 mmHg *from the participant's baseline

* The increase must be present at 2 measurements taken 6 hours apart

◆ First baseline or booking BP reading documented prior to 20+0 weeks gestation

BP reading /

◆ Highest BP Reading documented after 20+0 weeks AND prior to delivery time

BP reading / Highest systolic reading Highest diastolic reading

◆ Dipstick: Please record baseline protein on dipstick documented on antenatal records

☐ Negative ☐ Trace ☐ +1 ☐ +2 ☐ +3 ☐ +4

◆ Dipstick: Please record highest protein on dipstick documented in chart

☐ Negative ☐ Trace ☐ +1 ☐ +2 ☐ +3 ☐ +4

◆ Was a 24 hr Urine done? ☐ Yes ☐ No 24 hr Urine result . g/day

◆ If Hypertension in Pregnancy, Please check all that apply ☐ NONE

☐ Systolic BP $\geq$ 160 mmHg OR diastolic BP $\geq$ 110 mmHg recorded on at least 2 occasions at least 6 hours apart

☐ Oliguria (500 ml or less in 24 hours) ☐ Thrombocytopenia g/L (Please record lowest platelet count recorded)

☐ Cerebral or visual disturbances ☐ Urate umol/L

☐ Epigastric pain ☐ Pulmonary edema or cyanosis

☐ Increase in AST(>31U/L) U/L ☐ Oligohydramnios

☐ Increase in ALT (>40U/L) U/L ☐ MgSO4 therapy during intrapartum period

☐ Elevated serum creatinine (>75 umol/L) umol/L ☐ Labetolol IV during intrapartum period

☐ The occurrence of seizures

◆ Does this subject meet the criteria for HELLP? ☐ Yes ☐ No

HELLP syndrome is a group of symptoms that occur in pregnant women who have:

H- Hemolysis

EL- Elevated liver enzymes

LP- Low platelets count (<100,000 10⁹ /L)

Appendix 3: Kappa calculations for up to three raters and three categories

In our study we considered the simplified case of two raters and two categories, and then we generalized for the case of three raters and multiple categories. We are not considering cases of more than three raters since the tables and the formulae will be too complicated and out of the use in our specific case.

Let us consider the two categories are Positive and Negative. Now we can construct the 2 X 2 table of the two raters', also called readers and observers, joint agreement. As shown in Table A3.1.

As a convention we will assign capital letters to observed values and small letters for expected values. Two general indices of agreement can be derived from table A3.1, which are 'A' and 'D' and their expected values, 'a' and 'd', are calculated as shown in the table. The overall proportion of agreement, or observed proportion, is denoted by P_o , and

In case of 'K' categories and two raters, the joint agreement table take the following appearance (Table A3.2).

$$P_o = \frac{A + D}{n}$$

The expected proportion P_e can be calculated as follows:

$$P_e = \frac{a + d}{n}$$

And Kappa as

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

Table A3.1 Agreement calculations for 2 categories and 2 raters

		Reader 1		
		Positive	Negative	Total
Reader 2	Positive	A $a = G1 * F1 / n$	B $b = G1 * F2 / n$	$G1 = A + B$
	Negative	C $c = G2 * F1 / n$	D $d = G2 * F2 / n$	$G2 = C + D$
	Total	$F1 = A + C$	$F2 = B + D$	n

Table A3.2 Agreement calculations for “K” categories for 2 raters

		Reader 1			
		j = 1		j = K	Total
Reader 2	j = 1	O_{11} $e_{11} = G_1 * F_1 / n$		O_{K1} $e_{K1} = G_1 * F_K / n$	$G_1 = O_{11} + O_{21} + \dots + O_{K1}$ $G_K = \sum_{j=1}^K O_{j1}$
					
	j = K	O_{1K} $e_{1K} = G_K * F_1 / n$		O_{KK} $e_{KK} = G_K * F_K / n$	$G_K = \sum_{j=1}^K O_{j1}$
	Tot al	$F_1 = \sum_{j=1}^K O_{1j}$		$F_K = \sum_{j=1}^K O_{Kj}$	n

And P_o become

$$P_o = \frac{\sum_{j=1}^K O_{jj}}{n}$$

And P_e become

$$P_e = \frac{\sum_{j=1}^K e_{jj}}{n}$$

Where O_{jj} and e_{jj} are the observed and expected agreements respectively for category j and $j = 1, 2, \dots, K$. O_{jj} and e_{jj} are the diagonal values in the $K \times K$ table of two readers joint agreement. The expected values e_{jj} are calculated as follows

$$e_{jj} = \frac{F_j G_j}{n}$$

In case of ‘ K ’ categories and ‘ R ’ raters, the table will be more complicated since it will have an R dimensions, ie $K \times K \times K \dots \times K$ or K^R . However, we still are able to write P_o and P_e and P_o become

$$P_o = \frac{\sum_{j=1}^K O_{j\dots j}}{n}$$

And P_e become

$$P_e = \frac{\sum_{j=1}^K e_{j\dots j}}{n}$$

Where $O_{j\dots j}$ and $e_{j\dots j}$ are the observed and expected agreements respectively for category j and $j = 1, 2, \dots, K$. The notation $j\dots j$ mean j repeated R times. For Example if $R = 5$ then $j\dots j$ means $jjjjj$. $O_{j\dots j}$ and $e_{j\dots j}$ are the diagonal values in the K^R table of ‘ R ’ readers joint agreement. The expected values $e_{j\dots j}$ are calculated as follows

$$e_{j\dots j} = \frac{r_{1j} * r_{2j} * \dots * r_{Rj}}{n^{R-1}}$$

More conveniently, $e_{j\dots j}$ can be written as

$$e_{j,...,j} = \frac{\prod_{i=1}^R r_{ij}}{n^{R-1}}$$

Where r_{ij} is the total number of patients or subjects that rater i diagnosed them to have category j . In case of two raters and two categories $F_1 = R_{11}$, $F_2 = R_{12}$, $G_1 = R_{21}$, $G_2 = R_{22}$.

We are going to limit our case to three categories and three raters. The table is drawn as a 3 X 9 instead of 3 X 3 X 3 to fit with the two dimensional writing papers (Table A3.3).

From the table, $P_o = 176/205 = 0.86$ and $P_e = 0.24$ and as a result $Kappa = 0.82$ or 82%.

We have eliminated the subjects to whom at least one rater did not assign a category and constructed Table A3.4. From the table $P_o = 176/202 = 0.87$ and $P_e = 0.25$ and as a result $Kappa = 0.83$ or 83%.

Table A3.3 Agreement calculations for 3 categories and 3 raters

				READER 1			
				1	2	3	Total
R E A D E R 3	1	R		$O_{111}=76$	$O_{211}=2$	$O_{311}=0$	
		E	1	$e_{111}=16.20$	e_{211}	e_{311}	$V_1=O_{111}+O_{211}$
		A		$=r_{11}*r_{21}*r_{31}/n^2$	$=r_{12}*r_{21}*r_{31}/n^2$	$=r_{13}*r_{21}*r_{31}/n^2$	$+O_{311}=78$
		D		$O_{121}=4$	$O_{221}=3$	$O_{321}=0$	$V_2=O_{121}+O_{221}$
	2	E	2	e_{121}	e_{221}	e_{311}	$+O_{321}=7$
		R		$=r_{11}*r_{22}*r_{31}/n^2$	$=r_{12}*r_{22}*r_{31}/n^2$	$=r_{13}*r_{21}*r_{31}/n^2$	
			3	$O_{131}=0$	$O_{231}=0$	$O_{331}=0$	$V_3=O_{131}+O_{231}$
				e_{131}	E_{231}	E_{331}	$+O_{331}=0$
				$=r_{11}*r_{23}*r_{31}/n^2$	$=r_{12}*r_{23}*r_{31}/n^2$	$=r_{13}*r_{23}*r_{31}/n^2$	
		Total		$r_{31}=O_{111}+O_{121}+O_{131}+O_{211}+O_{221}+O_{231}+O_{311}+O_{321}+O_{331}=85$			
R E A D E R 3	2	R		$O_{112}=7$	$O_{212}=2$	$O_{312}=0$	
		E	1	e_{112}	e_{212}	e_{312}	$V_4=O_{112}+O_{212}$
		A		$=r_{11}*r_{21}*r_{31}/n^2$	$=r_{12}*r_{21}*r_{31}/n^2$	$=r_{13}*r_{21}*r_{31}/n^2$	$+O_{312}=9$
		D		$O_{122}=2$	$O_{222}=99$	$O_{322}=1$	$V_5=O_{122}+O_{222}$
		E	2	e_{122}	$e_{222}=33.13$	e_{312}	$+O_{322}=102$
		R		$=r_{11}*r_{22}*r_{32}/n^2$	$=r_{12}*r_{22}*r_{32}/n^2$	$=r_{13}*r_{21}*r_{32}/n^2$	
R E A D E R 3	2		3	$O_{132}=0$	$O_{232}=0$	$O_{332}=2$	$V_6=O_{132}+O_{232}$
				e_{132}	E_{232}	E_{332}	$+O_{332}=2$
				$=r_{11}*r_{23}*r_{32}/n^2$	$=r_{12}*r_{23}*r_{32}/n^2$	$=r_{13}*r_{23}*r_{32}/n^2$	

			Total	$r_{32} = O_{112} + O_{122} + O_{132} + O_{212} + O_{222} + O_{232} + O_{312} + O_{322} + O_{332} = 113$			
3	R E A D E R 2	1	$O_{113} = 0$	$O_{213} = 0$	$O_{313} = 0$	$V_7 = O_{113} + O_{213}$ $+ O_{313} = 0$	
			e_{113}	e_{213}	e_{313}		
		$= r_{11} * r_{21} * r_{33} / n^2$	$= r_{12} * r_{21} * r_{33} / n^2$	$= r_{13} * r_{21} * r_{33} / n^2$			
		2	$O_{123} = 0$	$O_{223} = 2$	$O_{323} = 0$	$V_8 = O_{123} + O_{223}$ $+ O_{323} = 2$	
			e_{123}	e_{223}	e_{323}		
		$= r_{11} * r_{22} * r_{33} / n^2$	$= r_{12} * r_{22} * r_{33} / n^2$	$= r_{13} * r_{21} * r_{33} / n^2$			
3	$O_{133} = 1$	$O_{233} = 3$	$O_{333} = 1$	$V_9 = O_{133} + O_{233}$ $+ O_{333} = 5$			
	e_{133}	e_{233}	$e_{333} = 0.005$				
			$= r_{11} * r_{23} * r_{33} / n^2$	$= r_{12} * r_{23} * r_{33} / n^2$	$= r_{13} * r_{23} * r_{33} / n^2$		
		Total	$r_{33} = O_{113} + O_{123} + O_{133} + O_{213} + O_{223} + O_{233} + O_{313} + O_{323} + O_{333} = 7$				
		Total	$r_{21} = V_1 + V_4 + V_7 = 89$		$r_{22} = V_2 + V_5 + V_8 = 111$	$r_{23} = V_3 + V_6 + V_9 = 7$	
		Total	$r_{11} = O_{111} + O_{121}$ $+ O_{131} + O_{112} +$ $O_{122} + O_{132} + O_{113}$ $+ O_{123} + O_{133}$ $= 90$	$r_{12} = O_{211} + O_{221}$ $+ O_{231} + O_{212} +$ $O_{222} + O_{232} + O_{213}$ $+ O_{223} + O_{233}$ $= 111$	$r_{13} = O_{311} + O_{321}$ $+ O_{331} + O_{312} +$ $O_{322} + O_{332} + O_{313}$ $+ O_{323} + O_{333}$ $= 4$	$n = 205$	

Table A3.4 Agreement calculations for 3 categories for 3 raters (without the subjects to whom at least one rater did not assign a category)

				READER 1			
				1	2	3	Total
R E A D E R	1	R					
		E	1	$O_{111} = 76$ $e_{111} = 16.13$ $= r_{11} * r_{21} * r_{31} / n^2$	$O_{211} = 2$ e_{211} $= r_{12} * r_{21} * r_{31} / n^2$	$O_{311} = 0$ e_{311} $= r_{13} * r_{21} * r_{31} / n^2$	$V_1 = O_{111} + O_{211}$ $+ O_{311} = 78$
		D					
		E	2	$O_{121} = 3$ e_{121} $= r_{11} * r_{22} * r_{31} / n^2$	$O_{221} = 4$ e_{221} $= r_{12} * r_{22} * r_{31} / n^2$	$O_{321} = 0$ e_{311} $= r_{13} * r_{21} * r_{31} / n^2$	$V_2 = O_{121} + O_{221}$ $+ O_{321} = 7$
	2	R					
		E	3	$O_{131} = 0$ e_{131} $= r_{11} * r_{23} * r_{31} / n^2$	$O_{231} = 0$ E_{231} $= r_{12} * r_{23} * r_{31} / n^2$	$O_{331} = 0$ E_{331} $= r_{13} * r_{23} * r_{31} / n^2$	$V_3 = O_{131} + O_{231}$ $+ O_{331} = 0$
		Total		$r_{31} = O_{111} + O_{121} + O_{131} + O_{211} + O_{221} + O_{231} + O_{311} + O_{321} + O_{331} = 85$			
	2	R					
		E	1	$O_{112} = 7$ e_{112} $= r_{11} * r_{21} * r_{31} / n^2$	$O_{212} = 2$ e_{212} $= r_{12} * r_{21} * r_{31} / n^2$	$O_{312} = 0$ e_{312} $= r_{13} * r_{21} * r_{31} / n^2$	$V_4 = O_{112} + O_{212}$ $+ O_{312} = 9$
		D					
		E	2	$O_{122} = 2$ e_{122} $= r_{11} * r_{22} * r_{32} / n^2$	$O_{222} = 100$ $e_{222} = 34.42$ $= r_{12} * r_{22} * r_{32} / n^2$	$O_{322} = 1$ e_{312} $= r_{13} * r_{21} * r_{32} / n^2$	$V_5 = O_{122} + O_{222}$ $+ O_{322} = 103$
	2	R					
		E	3	$O_{132} = 0$ e_{132}	$O_{232} = 0$ E_{232}	$O_{332} = 2$ E_{332}	$V_6 = O_{132} + O_{232}$ $+ O_{332} = 2$

3				$= r_{11} * r_{23} * r_{32} / n^2$	$= r_{12} * r_{23} * r_{32} / n^2$	$= r_{13} * r_{23} * r_{32} / n^2$	
			Total	$r_{32} = O_{112} + O_{122} + O_{132} + O_{212} + O_{222} + O_{232} + O_{312} + O_{322} + O_{332} = 114$			
	3	R E A D E R 2	1	$O_{113} = 0$ e_{113} $= r_{11} * r_{21} * r_{33} / n^2$	$O_{213} = 0$ e_{213} $= r_{12} * r_{21} * r_{33} / n^2$	$O_{313} = 0$ e_{313} $= r_{13} * r_{21} * r_{33} / n^2$	$V_7 = O_{113} + O_{213}$ $+ O_{313} = 0$
			2	$O_{123} = 0$ e_{123} $= r_{11} * r_{22} * r_{33} / n^2$	$O_{223} = 2$ e_{223} $= r_{12} * r_{22} * r_{33} / n^2$	$O_{323} = 0$ e_{323} $= r_{13} * r_{21} * r_{33} / n^2$	$V_8 = O_{123} + O_{223}$ $+ O_{323} = 2$
			3	$O_{133} = 1$ e_{133} $= r_{11} * r_{23} * r_{33} / n^2$	$O_{233} = 0$ e_{233} $= r_{12} * r_{23} * r_{33} / n^2$	$O_{333} = 0$ $e_{333} = 6.6 \times 10^{-4}$ $= r_{13} * r_{23} * r_{33} / n^2$	$V_9 = O_{133} + O_{233}$ $+ O_{333} = 1$
			Total	$r_{33} = O_{113} + O_{123} + O_{133} + O_{213} + O_{223} + O_{233} + O_{313} + O_{323} + O_{333} = 3$			
			Total	$r_{21} = V_1 + V_4 + V_7 = 87$		$r_{22} = V_2 + V_5 + V_8 = 112$	$r_{23} = V_3 + V_6 + V_9 = 3$
			Total	$r_{11} = O_{111} + O_{121}$ $+ O_{131} + O_{112} +$ $O_{122} + O_{132} + O_{113}$ $+ O_{123} + O_{133}$ $= 89$	$r_{12} = O_{211} + O_{221}$ $+ O_{231} + O_{212} +$ $O_{222} + O_{232} + O_{213}$ $+ O_{223} + O_{233}$ $= 110$	$r_{13} = O_{311} + O_{321}$ $+ O_{331} + O_{312} +$ $O_{322} + O_{332} + O_{313}$ $+ O_{323} + O_{333}$ $= 3$	n = 202

Appendix 4

Sample size calculations

Because disease prevalence is <0.50 and equal 5%, the sample size calculation has been done following this guideline: the first step requires an assumption on the expected value of the new diagnostic test sensitivity which is 70% according to the literature; the second step is to specify the minimum acceptable lower confidence limit (0.55 in our example), together with the required probability (which was set at 0.95) that this limit is not violated. The minimal sample size for the group of cases is then read from the (Table A4.1). The corresponding number of controls is obtained from equation:

$$N_{\text{controls}} = \frac{N_{\text{cases}} [(1 - \text{Prev})/\text{Prev}]}{0.95} = 114 [(1-0.05)/0.05] = 2736$$

The probability that the estimated 95% lower confidence limit is above the minimal acceptable value is 0.95.

Table A4.1 Number of cases (or controls) for expected sensitivities (or specificities)

Expected sensitivity (or specificity)	Minimal acceptable lower confidence limit											
0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95			
0.6				268	1058							
0.65				119	262	1018						
0.7				67	114	248	960					
0.75				42	62	107	230	869				
0.8				28	40	60	98	204	756			
0.85				18	26	33	52	85	176	624		
0.9				13	18	24	31	41	70	235	474	
0.95				11	12	14	16	24	34	50	93	298

Appendix 5 Ethics Approval



Ottawa Hospital Research Ethics Boards / Conseils d'éthique en recherches

751 Parkdale Avenue Suite 106, Ottawa, Ontario K1Y 1J7 613-798-5555 ext. 14902 Fax: 613-761-4311
<http://www.ohri.ca/ohreb>

Thursday, January 21, 2010

[REDACTED]
Ottawa Hospital - General Campus
Obstetrics & Gynecology
501 Smyth Road
Ottawa, ON
K1H 8L6

Dear [REDACTED]

Re: Protocol # 2010005-01H Validity of Administrative Database for Reporting Pre-eclampsia

Protocol approval valid until - Thursday, January 20, 2011

Thank you for your email dated January 12, 2010. I am pleased to inform you that your Application for Chart Review underwent expedited review by the Ottawa Hospital Research Ethics Board (OHREB), and is approved. No changes, amendments or addenda may be made to the protocol without the OHREB's review and approval.

If the study is to continue beyond the expiry date noted above, a Renewal Form should be submitted to the OHREB approximately six weeks prior to the current expiry date. If the study has been completed by this date, a Termination Report should be submitted.

The Ottawa Hospital Research Ethics Board is constituted in accordance with, and operates in compliance with the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans; Health Canada Good Clinical Practice: Consolidated Guideline; Part C Division 5 of the Food and Drug Regulations of Health Canada; and the provisions of the Ontario Health Information Protection Act 2004 and its applicable Regulations.

Yours sincerely,

[REDACTED]

Chairman
Ottawa Hospital Research Ethics Board