9-9-2010

Validity of self-report with respect to prescription medications among pregnant women

Pallavi Jaiswal

Pallavi Jaiswal
Candidate

College of Pharmacy- Pharmaceutical Sciences
Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Ludmila Bakhireva ,Chairperson

Nasreen Khan

Melanie Dodd

VALIDITY OF SELF-REPORT WITH RESPECT TO PRESCRIPTION MEDICATIONS AMONG PREGNANT WOMEN

BY

PALLAVI JAISWAL

B.PHARM., 2007
INDRAPRASTHA UNIVERSITY

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Pharmaceutical Sciences

The University of New Mexico
Albuquerque, New Mexico

July, 2010

DEDICATION

I dedicate this thesis to my grandparents, Late Sh. Satyadev Chaudhary, Late Sh. Moti Lal Jaiswal, Late Mrs. Sona Jaiswal, and Mrs. Draupadi Devi; my beloved parents Vivekanand and Neelam; and my dearest sister Rashi. I owe everything I have achieved to you all.

ACKNOWLEDGMENT

I would like to express my sincere gratitude to my advisor and thesis committee chair, Dr. Ludmila Bakhireva. She has been a constant source of support and encouragement during the course of this thesis project. Her tremendous energy, knowledge, and commitment for research have inspired me. She has always made time to answers all my questions, some important, some trivial, and she has always done so with a smile. I owe her special thanks to help me write my thesis and her invaluable suggestions that helped me clarify and better articulate my ideas. For that and for everything else she has done along the way, I will always be grateful.

I would also like to thank my committee members, Dr. Nasreen Khan and Dr. Melanie Dodd for their insightful comments and reviewing my work. I would like to especially thank Dr. Nasreen Khan for teaching me how to conduct quality research and the time she has dedicated to help me in understanding the fundamentals.

The support of my friends during the course of my graduate studies was something I will never forget. They have always made me smile whenever I needed it. Thank you all for being so encouraging and cheerful!

Words cannot express the support and love I have received from my family. My parents, Neelam and Vivekanand have always believed in me and have been my greatest supporters in whatever I chose to so. Their advice, encouragement and support have always been my greatest strength. I would not have achieved what I have without their love and guidance. I would like to thank my sister Rashi for being so patient and listening to my complaints and frustrations and always cheering me up whenever I was down. Munni mausi, Guddu mausi, Gullu mama, and Deepak mama- you all have always been so encouraging. I am the luckiest person to have you all by my side.

VALIDITY OF SELF-REPORT WITH RESPECT TO PRESCRIPTION MEDICATIONS AMONG PREGNANT WOMEN

BY

PALLAVI JAISWAL

ABSTRACT OF THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Pharmaceutical Science

The University of New Mexico
Albuquerque, New Mexico

July, 2010

VALIDITY OF SELF-REPORT WITH RESPECT TO PRESCRIPTION MEDICATIONS AMONG PREGNANT WOMEN

By

Pallavi Jaiswal

B.Pharm., Indraprastha University, 2007

M.S., Pharmaceutical Sciences, 2010

ABSTRACT

Objectives: Accuracy of self-report regarding prescription medication use among pregnant women is largely unknown. Accurate self-reported information is needed for medication reconciliation purposes, clinical management, clinical teratology research, and monitoring of adherence. This study examined the accuracy of self-reported medication use by pregnant women for medications used chronically and episodically or intermittently during pregnancy. Further, predictors of inaccurate reporting regarding prescription medication use were also estimated.

Methods: A cross-sectional analysis of data collected through the University of New Mexico (UNM) cohort study, "Safety of Medication and Perception of Teratogenicity" (SMART) was conducted. Pregnant women were recruited from UNM prenatal care clinics and were asked to report all medications they took since their last menstrual period. The analysis was limited to women enrolled in the first year of the study who had at least one prescription for diabetes or opioid analgesics medications (representative of chronic and acute medication use, respectively). The accuracy of agreement between self-

report and medical records for each medication class was estimated by simple (κ) and prevalence and bias adjusted (PABAK) kappa. Information from the medical records was used as the 'gold-standard'. Multivariable logistic regression analyses were used to determine predictors of inaccurate reporting of prescription medication use in this cohort.

Results: A total of 92 pregnant women were included in the analysis. Agreement for diabetes medications was near perfect (κ=0.87; PABAK=0.91); whereas poor-to-moderate concordance was observed for opioid analgesics (κ=0.29; PABAK=0.57). Among antidiabetic medications, concordance was highest for biguanides (κ=0.90; PABAK=0.93) and lowest for sulfonylureas (κ=0.83; PABAK=0.87); whereas among opioid analgesics, highest agreement was observed for strong agonists (κ=0.51; PABAK=0.56) and lowest for moderate/low agonists (κ=0.06; PABAK=0.59).Women engaging in at least one episode of binge drinking were found to be inaccurate reporters of medication regarding prescription medication use (OR: 3.40, 95% CI: 1.13;10.29).

Conclusions: This study suggests poor accuracy of self-report with respect to prescription medications used as short courses or intermittently during pregnancy. Therefore, in clinical studies assessing safety of such medications in pregnancy, self-reported information needs to be supplemented by other sources. Accuracy of self-report for medications used chronically is acceptable.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER – 1 INTRODUCTION

1.1 Prescription medication use during pregnancy

Prescription medication use during pregnancy in the United States is fairly high [1, 2]. Estimates put the proportion of pregnant women taking at least one prescription medication (excluding vitamin and iron preparations) during pregnancy in the range 56-80% [2, 3], and approximately 32% pregnant women on an average use two prescription medications during pregnancy [2, 4, 5].

Researchers have indicated that about 40-60% pregnant women are prescribed prescription medications with unknown safety profile during pregnancy [2, 6]. Also, given the high rate of unintended pregnancies (49%) in the U.S. [7], many women might have accidental exposure in the early weeks of pregnancy to potentially teratogenic medications before realizing that they are pregnant [7]. Such high rates are therefore alarming when one considers the risk of accidental exposure to these medications and the associated risks [7]. Hence, during pregnancy it is vital that accurate information about medications and their risks is made accessible to both the patient and the physician.

In the United States, the Food and Drug Administration (FDA) has established a classification system that assigns a letter category to medications on the basis of potential to cause birth defects/teratogenic effects, ranging from class A (classified as being relatively 'safe' for use during pregnancy), to class X (classified as being contraindicated for use during pregnancy, because evidence of potential risk posed by the medication outweigh the possible benefits) [8]. Table 1 lists the above mentioned risk classification system.

Table 1: The U.S. Food and Drug Administration (FDA) teratogenic risk categories [8]

Category	Description
A	Controlled studies in women fail to demonstrate a risk to the fetus in the first trimester; possibility of fetal harm appears remote.
B	Either animal studies do not indicate a risk to the fetus, and there are no controlled studies in women; or animal studies have shown an adverse effect, but controlled studies in women failed to demonstrate a risk.
C	Either animal studies indicate a fetal risk, and there are no controlled studies in women; or studies in women and animals are not available.
D	There is positive evidence of fetal risk, but the benefits may be acceptable despite the risk.
X	There is definite risk based on studies in animals or humans or based on human experience, and the risk clearly outweighs any possible benefit.

While, this classification provided by the FDA is widely used in interpreting the risk of teratogenicity associated with prescription medications used during pregnancy, it is also criticized as being ambiguous and misleading that can create confusion and concern among pregnant women regarding medication use during pregnancy [1, 9]. Recently, FDA has proposed to make extensive changes to the existing classification system to make it more clear and understandable [10, 11]. As a part of the proposed change, the exiting system of classification would be eliminated. The new pregnancy labeling for medications would contain three sections: risk summary, clinical considerations and data section [10, 11]. The risk summary section would classify the

likelihood of the drug causing various birth defects [10, 11]. Clinical considerations section would contain detailed information about prescribing decisions for pregnant women, dosing adjustments, and also information about inadvertent exposure to the medication [10, 11]. Finally, the data section would describe whether the information is from animal or human data, and will also describe that data in detail [10, 11]. The proposed changes would allow a broader spectrum of data, such as pregnancy registry data, to be included in the determination of potential maternal and fetal harm [10].

1.2 Self-reported information: importance and problems

Data collection through self-reports is the most common method of gathering information in epidemiologic studies or surveys [12-17]. Self-reported information can be collected through self-administered questionnaires, face-to-face interviews, or telephone interviews. Information collected by this method can range from information about general health status of the subjects, information regarding specific disease conditions, behavioral and lifestyle characteristics, to information regarding prescription medication use.

Information collected by self-report plays an essential part in research to estimate prevalence of a disease, access to health care, health care delivery, preventive behaviors and utilization of health care services [18]. Various surveys are conducted at the national level, e.g., Behavioral Risk Factor Surveillance System (BRFSS), National Health and Nutrition Examination Survey (NHANES), and National Health Interview Survey (NHIS). Most of these national surveys, like NHANES, BRFSS, The National Hospital

Ambulatory Medical Care Survey (NHAMCS), also contain information about medication use from the respondents. Information derived from these surveys has been helpful in research, public health planning, and developing polices, health/preventive campaigns, and improving health care delivery and access to care [19].

Physicians also routinely obtain information from the patients regarding their medication use and often rely on patient self-report [5]. Information about medication use is a vital component for the success and completion of a treatment regime. This information is essential for medication reconciliation purposes and also to individualize medication therapy suited to the patients' needs [20]. Medication reconciliation, as defined by the Institute of Healthcare Improvement (IHI), is the process of reviewing and comparing a patient's complete medication profile and history [21]. It is performed in order to provide accurate and correct medication to the patients, and to prevent discrepancies in medical care by identifying sources of potential medication errors, such as duplications, omissions, dosing errors, adverse drug reactions and potential drug interactions [21, 22]. Examples of medication errors include misreading or miswriting a prescription. An adverse drug reaction occurs in response to a drug which is noxious and unintended and which occurs at doses normally used in man for prophylaxis, diagnosis, or therapy of disease or for the modification of physiologic function [22]. Medication errors that are stopped before harm can occur are called potential adverse drug event [22]. Such errors are usually found to occur at various points during medical care, such as, at the time of admission to any health care facility, transfer within the facility, and at the time of discharge [22].

In the year 2005, The Joint Commission (formerly, the Joint Commission on Accreditation of Healthcare Organizations; (JCAHO)) had announced medication reconciliation to be one of the National Patient Safety Goals [21]. Even though medication reconciliation is considered an important exercise in the process of providing medical care, there is no standard procedure to perform it. Mostly, a multidisciplinary approach, consisting of participation from physicians, nurses, and pharmacists is utilized [23]. During the course of this study (i.e., in the year 2009), TJC required implementation of a process for obtaining and documenting a complete list of the patient's current medications upon the patient's admission to the organization and with the involvement of the patient [23].

During pregnancy, medication safety is vital in the light of issues related to the risk of teratogenicity. According to previous research, approximately 6.0-9.5% pregnant women in the United States are at risk of exposure to medications with potential teratogenic effects [7, 24]. Lo et al. also found that over a period of 20 years (1980-2000), information regarding the safety of about 90% of the newly marketed drugs approved by the FDA was insufficient with regard to safety profile in pregnancy [24].

Often times during pregnancy, women might have some illness (pre-existing or developed during pregnancy), such as diabetes mellitus, hypertension, or other chronic illnesses that might necessitate drug treatment [9, 24]. Poor control of maternal conditions increases the risk of birth defects [2, 9, 25]. Moreover, sometimes women misinterpret recommendations about the use of prescription medications during pregnancy. For example, they might misinterpret recommendation that the use of a

particular medication during pregnancy is not recommended due to potential safety concerns [9]. It has also been documented that due to the lack of information about teratogenic risk posed by some medications, women can overestimate the risk of teratogenic effects [24]. As a result, such perception can result in unnecessary anxiety or even a termination of pregnancy [24]. Thus, medication reconciliation in pregnant women is of particular importance, given the potential teratogenic effect of some medications and the need to optimize therapeutic management of underlying maternal condition/s during pregnancy [24].

1.3 Self-report bias

It has been reported that several limitations are associated with self-reported information. These can potentially arise from various information biases, like recall bias, interviewer bias, and social desirability bias [13, 16, 17, 26-29]. *Recall bias* occurs when the response of the respondents in a survey or interview is affected by limits of memory. Moreover in retrospective studies, recall among study subjects with adverse outcomes (cases) can be substantially different from recall of healthy participants (controls). This would also result in recall bias. *Interviewer bias* takes place when the survey instrument is not self-administered, in which case the researcher is familiar with the study hypothesis that further can potentially influence the responses and their interpretation. *Social desirability bias* is another type of response bias when the respondents intentionally give responses in a socially acceptable or desirable manner. Such type of bias is mostly seen when the respondents are asked sensitive questions which they might deem inappropriate or uncomfortable to answer.

In the elderly population and in cohorts of patients with chronic illnesses, it has been found that such response biases occur mostly due to limited memory such as, forgetting the diagnosis, and presence of other comorbid conditions. In the case of inaccurate reporting by pregnant women, it has been found that recall bias occurs because pregnant women do not find episodic or acute events to be significant enough to report to the physician as compared to chronic illnesses/conditions, or they might not be able to recall them because they are already taking medication for other conditions [5, 30]. Maternal recall in this regard has also been found to be affected by history of adverse perinatal outcome, with events and exposures that took place in a pregnancy associated with any adverse perinatal outcomes being recalled more accurately than those events that occurred in a normal pregnancy [3, 31]

There is also evidence of social desirability bias in self-reported information provided by pregnant women. It has been documented that pregnant women are unwilling to disclose information about certain conditions or medications because of social stigma associated with those conditions or behaviors. This is seen commonly in situations when pregnant women are taking medication for conditions such as depression, sexually-transmitted disease or substance abuse; all of which can be perceived as negative in society [16, 29]. Similar evidence is present for alcohol consumption during pregnancy. Women have discomfort in revealing information about alcohol consumption during pregnancy because of fear of social stigma. Such behavior is deemed inappropriate as it can harm the developing fetus. Researchers have also indicated that because of rising awareness about these risky behaviors during pregnancy (smoking, drug use, and alcohol consumption), women might feel embarrassed while disclosing such information, which

might suggest that they are not able to curb their addiction to such risky behaviors even during pregnancy [32-35]. Lastly, with regard to prescription medication use pregnant women might not want to disclose information regarding consumption of medications belonging to a high risk category (Category D or X), as it might be considered negligent behavior with respect to the safety of the fetus [25], especially in cases of inadvertent exposure. Thus, self-reported information cannot be relied upon solely to make clinical decisions, and ascertainment of the accuracy of such information is vital in the light of drawbacks that it presents.

1.4 The concept of 'gold-standard' with respect to accuracy of medication use

Usually, the accuracy of self-reported data is determined by comparing it with some other source of information, considered to be the 'gold-standard'. Often, information recorded in patients' medical records is considered as the 'gold standard'. However, one of the major limitations of such 'gold-standard' is that it often does not contain information about medications prescribed outside the clinic/hospital where the patient is currently seeking care [36-38]. In addition, other limitations associated with medical records include delayed and erratic recording, illegible handwriting of physicians, non-documentation of medications taken from an outside provider, and very limited information about lifestyle and behavioral characteristics [26, 36, 39]. All of these can influence the quality of the information abstracted from the medical records.

Medical records are a good source of information for medical/illness history of a

patient, since in most of the cases they are accompanied by results of diagnostic procedures confirming the diagnosis of the condition. However, the same cannot be said regarding information recorded in medical records about medication use, especially in cases where this information is documented on the basis of self-report by the patient. Such information can be misleading particularly in the cases of medications used for the treatment of acute or episodic illnesses. This can be attributed to the fact that acute illnesses do not continue for a long period of time. On the contrary, most chronic illnesses have a well-defined diagnostic criterion, continue for a prolonged period of time and often require regular medication use.

Alternatively, other sources of information can be used for ascertaining accuracy of self-reported information, e.g., information on prescription fills obtained from pharmacy databases [16, 28, 29], administrative prescription registries [40], and wherever possible, biochemical tests, e.g., like serum/urine cotinine levels for estimating smoking exposure [32, 33, 35, 41, 42] . Medical records and prescription databases are reported to be the most reliable sources [43].

1.5 Validity of self-reported information among general population with respect to chronic and acute conditions

In the general population, a number of studies have assessed the validity of self-reported information for a range of medical conditions and medication use [12-15, 18, 20, 36, 38, 44-53]. Agreement in previous studies has been found to vary by the study

population (e.g., elderly, cohorts of patients diagnosed with a specific condition, members of health plans having a specific medication use), specific disease/condition under investigation and the analytic method used to assess validity [15, 48]. However, despite these inconsistencies in the study population and study methodology, researchers have reported that patients provide reliable information about chronic conditions in comparison with acute conditions. Researchers attribute this differential recall to various reasons ranging from variability in diagnostic criteria and severity of the condition to frequency of patient-physician visits associated with some chronic conditions [12, 15, 26, 38, 54].

Similar evidence is present for accuracy of self-reported medication use. Typically, researchers have noted a higher accuracy of recall for medications used chronically in comparison with medications used as short-courses or episodically. This variation in recall of medication use has largely been ascribed to difference in the frequency and duration of medication use. In addition, researchers have also studied predictors that influence recall of medication use. Age, educational status and recall interval are the most commonly reported predictors.

1.6 Validity of self-reported information among pregnant women

While the accuracy of self-reported information regarding prescription medications in the general population has been validated in numerous studies, accuracy of self-reported information regarding prescription medication use among pregnant

women is largely unknown. In addition, there is dearth of information about recall accuracy among pregnant women regarding information about the prescription medications they take for chronic and acute conditions during their pregnancy.

Although knowledge of teratogenic potential is a critical part of a drug's benefit/risk profile, pregnant women are rarely included in clinical trial. There may be inadvertent pregnancy exposures during clinical trials of new products, but available data are usually insufficient to permit an adequately powered statistical analysis [25]. Consequently, when a drug is first marketed there are usually no human data on the effects of in-utero drug exposure. The only data on fetal effects initially available in the product labeling usually comes from animal reproductive toxicology studies [25]. Despite the lack of information on the safety of drug use during pregnancy, most pregnant woman likely would be exposed to drugs. Fetal exposure can occur before a woman knows she is pregnant. Some women enter pregnancy with medical conditions that require continuing drug therapy. New medical problems may develop during, or old ones may be exacerbated by, pregnancy.

Since clinical trials in pregnant women are unethical due to a potential harm from medication exposure to the developing fetus, observational studies are used to study the safety of medications in pregnancy [25]. Observational studies often rely on self-reported data or a combination of self-reported data with information from the medical or pharmacy records to make conclusions about teratogenicity or safety of specific medications.

1.7 Significance of the study and specific aims

Approximately 56-80% pregnant women use prescription medication during pregnancy [2, 3]. Most of these medications are used for reasons other than supplementing the nutritional requirements during pregnancy, e.g., treatment of an underlying or pre-existing condition. Diabetes is one such condition that affects approximately 3-10% of pregnancies in the United States, of which 90% is gestational diabetes and 8% is pre-existing [55]. Researchers also estimate that cases of gestational diabetes are rising at an annual rate of 8% [56]. Using data from a nationally representative sample of a cohort of pregnant women, Andrade et al. indicated that 7.9% of pregnant women in the United States are prescribed antidiabetic medication, of which 1.4% used insulin preparations [2]. In different European populations, researchers have estimated the prevalence of use of antidiabetic medications use to be in the range of 0.18-1.0% [43, 57, 58].

Analgesics are also prescribed frequently in pregnancy. In the above-mentioned study, Andrade et al. estimated that about 33.9% of pregnant women are prescribed opioid and nonopioid analgesics, of which 2.7% were prescribed codeine/guaifenesin preparations [2]. Earlier, Bracken et al. in their study noted that 0.4% of 1,427 pregnant women recruited from five hospitals in Connecticut were prescribed opioid analgesic medication [59]. In another study Piper et al. estimated that in a cohort of 18, 886 pregnant women receiving Medicaid, 25.6% were prescribed codeine containing preparations, and 4.0% were prescribed propoxyphene containing medications [4].

Further, Glover et al. reported that in a cohort of 578 pregnant women attending rural obstetric clinics in West Virginia, 5.0% reported use of prescription medications containing codeine, and 3.3% reported use of medications containing oxycodone [5]. Prevalence of specific opioid analgesic medications among pregnant women in the United States has been also been reported as 3.3% for oxycodone and 4% for propoxyphene [4, 5].

While these numbers suggest that a significant proportion of pregnant women do consume prescription medication during pregnancy, the key question is how many of them disclose information about their use. This study thus aims to evaluate the extent of agreement between self-reported information and information in the medical records regarding prescription medication use in a cohort of pregnant women enrolled in the "Safety of Medications During Pregnancy and Women's Perception of Teratogenic Risk (SMART) study" at the University of New Mexico (UNM), Albuquerque. In addition, it will also identify predictors of inaccurate reporting with respect to prescription medication use among these pregnant women.

The *central hypothesis* for this study is that the accuracy of self-reported information regarding prescription medications use among pregnant women is affected by the type of illness/condition (chronic vs. acute). This central hypothesis would be tested by the following specific aims:

Specific aim 1:

To compare the agreement between self-reported information and information in the

medical records in a cohort of pregnant women enrolled in the UNM-based SMART study cohort.

Research hypothesis 1:

Agreement between self-reported information and information in the medical records will be greater for medications prescribed for chronic medical conditions (represented by antidiabetic medication), compared with prescription medications given for acute medical conditions (represented by opioid analgesics).

Specific aim 2:

To examine the predictors of inaccurate reporting in this UNM-based study cohort.

Research hypothesis 2:

Certain maternal *demographic* (age, educational level, marital status, insurance coverage, ethnicity, place of birth, language), *lifestyle* (smoking status, alcohol use), *maternal medical and reproductive characteristics* (number of chronic conditions, gestational age, previous adverse perinatal outcomes, gravidity, parity) and participants' *knowledge and attitudes* towards medication use in pregnancy will influence the likelihood of disagreement between self-reported data and information in medical records.

The findings of the study would contribute to the limited knowledge that exists in the literature about the extent of agreement on prescription medication use between self report and medical records among pregnant women. Accurate self-reported information is needed for medication reconciliation purposes, clinical management, clinical teratology

research, and monitoring of adherence. Findings of this study would thus help healthcare providers and researchers to identify pregnant women who are more likely to inaccurately report information regarding their prescription medication use and also identify therapeutic classes which are especially prone to inaccurate reporting.

The purpose of collecting and evaluating data on drug exposure during pregnancy is to address whether a particular drug exposure increases the risk of abnormal fetal development above the background rate. With respect to research in the field of safety of medication in pregnancy, often self-report is the only method of obtaining information regarding drug exposure. Therefore, the findings of the study would contribute towards research on safety of medication use in pregnancy, which often relies on self-reported information to ascertain exposure. Findings of this study will also help in identifying the classes of medication which are at higher risk of being incorrectly reported by pregnant women. Relying on patient self-report alone for those classes can lead to incorrect estimation of medication exposure and erroneous assessment of teratogenicity.

CHAPTER-2 LITERATURE REVIEW

This chapter will provide an overview of the literature related to various studies that have assessed the extent of agreement and accuracy of recall between self-reported information and information from the medical records and/or other sources. The literature review is presented in two main sections: 1) assessment of agreement in the general population; 2) assessment of agreement in pregnant women.

2.1 Assessment of agreement in the general population

This section contains studies that have been conducted in various populations (non-pregnant) to assess the accuracy of recall and agreement. This section is divided into three parts: 1) assessment of agreement by condition; 2) assessment of agreement for prescription medication use; 3) predictors of recall and agreement. Table 2 presents a list of published kappa values of various medical conditions and medication use among general population.

Table 2: Accuracy of agreement for specific medical conditions and medication use: overview of literature

Medical conditions/medications	Kappa statistic (κ)	References
MEDICAL CONDITIONS		
Chronic conditions		
Diabetes	0.72 - 0.94	Kreigsman et al. [12], Haapanen et al. [13], Bush et al. [14], Okura et al. [15], Simpson et al. [20], Tisnado et

Medical conditions/medications	Kappa statistic (κ)	References
		al. [26], Skinner et al. [48], Merkin et al. [51], Iversen et al. [53], Leikauf et al. [60], Brownson et al. [61], Corser et al. [54], Martin et al. [62], Miller et al. [63].
Hypertension	0.24 – 0.85	Kreigsman et al. [12], Haapanen et al. [13], Okura et al. [15], Merkin et al. [51], Iversen et al. [53], Young et al. [64].
Asthma	0.43 – 0.78	Tisnado et al. [26], Iversen et al. [53], Corser et al. [54].
Musculoskeletal disorders	0.07 – 0.54	Kriegsman et al. [12], Simpson et al. [20], Skinner et al. [48], Goebeler et al. [49], Boissonnault et al. [52], Miller et al. [63]
Depression	0.11 – 0.40	Goebeler et al. [49], Leikauf et al. [60], Kwon et al. [65].
Congestive Heart Failure (CHF)	0.09 – 0.60	Okura et al. [15], Simpson et al. [20], Merkin et al. [51], Corser et al. [54], Miller et al. [63].
Myocardial Infarction (MI)	0.40 – 0.55	Merkin et al. [51], Young et al.[64].
Acute conditions		
Peripheral vascular disease	0.24 – 0.43	Kriegsman et al. [12], Simpson et al. [20], Corser et al. [54].
Lower back pain	0.33 – 0.54	Haapanen et al. [13], Skinner et al. [48], Miller et al. [63].
Pneumonia	0.27 – 0.62	Boissonnault et al., Iversen et al. [53]
Duodenal/peptic ulcer	0.14 – 0.29	Smith et al., Corser et al. [54]
Hay fever/rhinitis	0.40	Iversen et al. [53]

Medical conditions/medications	Kappa statistic (κ)	References
Shortness of breath	0.20	Tisnado et al [26].
MEDICATION USE		
Medications for treatment of chronic conditions		
Cardiovascular disease (CVD) medications	0.60 - 0.97	Tisnado et al. [26], Sjahid et al. [66], Brown et al. [67].
Specific classes of CVD medications		
Calcium channel blockers	0.91	Caskie et al. [68].
Beta-blockers	0.80 – 0.97	Caskie et al. [68], Sjahid et al. [66], Tisnado et al. [26].
ACE inhibitors	0.90	Caskie et al. [68].
Statins	0.88	Glintborg et al. [39]
Nitrates	0.20 – 0.31	Sjahid et al. [66], Tisnado et al. [26].
Cardiac glycosides	0.77 – 0.97	Caskie et al. [68], Glintborg et al. [39]
Vasodilating agents	0.73	Caskie et al. [68].
Hypotensive agents	0.83	Caskie et al. [68].
Antidepressant medications	0.42 – 0.77	Paganini-Hill et al. [16], Tisnado et al. [26], Caskie et al. [68], Haukka et al. [69], Kwon et al. [65].
HRT medication	0.21 – 0.92	Tisnado et al. [26], Nielsen et al. [70], Caskie et al. [68], Kropp et al. [71], Løkkegaard et al. [72].
Antidiabetic medications		
Specific classes		
Insulin	0.60 – 0.78	Tisnado et al. [26], Nielsen et al. [70]
Sulfonylureas	0.60 – 0.93	Glintborg et al. [39], Tisnado et al.[26].
Biguanides	0.60	Tisnado et al. [26].

Medical conditions/medications	Kappa statistic (κ)	References
TZDs	0.60	Tisnado et al.[26].
Medications for treatment of acute conditions		
Opioid analgesics	0.15 – 0.49	Caskie et al. [68], Nielsen et al. [70], Tisnado et al. [26].
NSAIDs	0.30 – 0.63	Tisnado et al. [26], Nielsen et al. [70], Caskie et al. [68], Kropp et al. [71], Løkkegaard et al. [72].
Gastrointestinal agents	0.50 – 0.67	Westbrook et al. [36], Nielsen et al. [70], Caskie et al. [68].
Musculoskeletal disorders	0.64 – 0. 96	Solomon et al. [50], Curtis et al. [73].

2.1.1 Assessment of agreement by condition

Studies conducted to evaluate agreement by condition, often focus on the ability of the respondents to recall diagnoses of chronic conditions. Self-reported information in such studies has been compared with a variety of sources considered to be the 'gold-standard', ranging from medical records and pharmacy records to physician notes. Accuracy of self-report in such cases is reported to vary by the study population, methodology, and the specific condition in question [15, 51, 60].

Studies have also reported difference in recall and agreement according to the diagnostic criteria of the condition. Researchers have consistently reported higher agreement for chronic conditions like diabetes and hypertension. Agreement for diabetes has been reported in the range $\kappa = 0.72$ to $\kappa = 0.94$ in different studies [12-15, 20, 26, 48,

51, 53, 54, 60-64], whereas for hypertension it has been reported in the range κ= 0.24 to κ= 0.85 [12, 13, 15, 51, 53, 64]. Studies have also reported agreement in the range 37%-73% for self-report of diabetes [17, 62]. Agreement for asthma has been reported in the range κ= 0.43 to κ= 0.78 [26, 53, 54]. Smith et al. reported 42% agreement for self-report of asthma [17]. Such high level of agreement for these conditions has been ascribed to the fact that these conditions have clear and well-defined diagnostic criteria that are easily understood by the patients, and hence easily recalled [15, 26, 38]. Continuous care and regular medication use is required to control these conditions that also helps in facilitating recall [15, 26, 38]. In addition, frequent contact with the healthcare providers upon diagnosis of such chronic conditions is also documented as one of the reasons that assists in recall [15, 26, 38, 54, 60, 64].

In contrast, low-to-moderate agreement has been reported for conditions like musculoskeletal disorders (κ=0.07 to κ=0.54) [12, 20, 48, 49, 52, 63], peripheral vascular disease (κ=0.24 to κ=0.43) [12, 20, 54], depression (κ=0.11 to κ=0.40) [49, 60], lower back pain (κ=0.33 to κ=0.54) [13, 48, 63], and claudication (impairment in walking) (κ=0.30) [13]. Low agreement has also been documented for specific cardiovascular diseases like, congestive heart failure (CHF) (κ=0.09 to κ=0.60) [15, 20, 51, 54, 63], myocardial infarction (MI) (κ=0.40 to κ=0.55) [51, 64]. The low level of agreement obtained for these conditions is largely attributed to the ambiguous and complex diagnostic criteria for these conditions which is not easily understood by patients [15, 26, 48, 54, 60, 63]. Conditions like CHF and peripheral vascular disease have irregular,

fluctuating and intermittent symptoms which further makes the recall of these conditions difficult [20, 54]. Subjective nature of conditions like lower back pain and arthritis is also reported as one of the reasons that attributes to lower agreement for these conditions, as researchers have reported that these conditions are more often reported by patients as compared to their documentation in the medical records [49, 54].

One of the main concerns expressed in the studies conducted to assess the accuracy of agreement by condition, has been the variability of recall because of difference by the condition (chronic vs. acute). Low-to-moderate agreement has been documented for pneumonia (κ=0.27 to κ=0.62) [52, 53], and duodenal/peptic ulcer (κ=0.14 to κ=0.29) [17, 54], hay fever/rhinitis (κ=0.40) [53], and shortness of breath (0.20) [26]. Episodic nature of these conditions has been attributed to the lower recall rates obtained for these conditions [17, 26, 38, 54].

Studies have also been conducted in cohorts of cancer patients to assess agreement between self-reported information and information from medical records. Their findings vary by report of type of screening test and cancer site. High agreement has been documented for self-report of breast cancer (91%), prostate and lung cancer (90% each), prostate-specific enzyme (PSA) test, sigmoidoscopy, and colonoscopy (κ=0.40 to κ=0.80) [45, 46]. One interesting finding by Mukerji et al. in their cohort of patients with neck and head cancer was that while reporting co-morbid conditions, patients reported some chronic conditions more accurately (diabetes, κ=0.89 and stroke, κ=0.77), than other (arthritis, κ=0.11) [47].

Summary:

In summary, the results of the studies conducted to assess agreement for self-report of various conditions have a high level of agreement for most chronic conditions compared with acute conditions or those with sporadic/intermittent symptoms. Even in the presence of other comorbid conditions, recall for a chronic condition was high. Accuracy of recall also is also affected by the variability in diagnostic criteria of medical conditions, which might lead the patient to misunderstand or misinterpret the diagnosis. Conditions with well-defined diagnostic criteria are recalled more easily by patients than those conditions with complex or ambiguous diagnostic criteria.

2.1.2 Assessment of agreement for prescription medication use

Studies conducted to assess agreement for self-reported prescription medication use often focus on therapeutic classes used for treatment of chronic conditions e.g., cardiovascular diseases (CVD), diabetes, musculoskeletal disorders, and hormone therapy (HRT).

In a study conducted in the Dutch population, Van den Brandt et al. found moderate agreement (61.2%) for recall of medications used in the past over a period of two years [29]. Moderate-to-high agreement has been recorded for CVD medication use across various studies (κ=0.60 to κ=0.97) [26, 66, 67]. Researchers have attributed this trend to the fact that these medications are taken regularly and for a long period of time, and are more likely to be accurately recalled [29, 70].

While overall agreement for self-reported use of any CVD medication was high in these studies, it varied considerably when comparison was made across specific drug classes. For example, Caskie et al. reported highest overall agreement for CVD drugs (κ=0.83) among the ten drug categories (including antihistamines, anti-infectives, autonomic drugs, blood formation and coagulation products, CVD drugs, electrolytic drugs, eye-ear-nose-throat preparations, gastrointestinal drugs, hormonal drugs) [68]. However, there was marked variation in the agreement across various drug classes for CVD medications (κ=0.91 for calcium channel blockers; κ=0.73 for vasodilating agents; κ=0.77 for cardiac glycosides; κ=0.84 for beta-blockers; κ=0.90 for ACE inhibitors; κ=0.83 for hypotensive agents) [68]. Glintborg et al. also reported high overall agreement for CVD drug use, but upon analyzing agreement by drug class, they found higher agreement for digoxin (κ=0.97) and slightly lower agreement for statins (κ=0.88) [39].

Similarly, Sjahid et al. reported high agreement for β-blockers in their study (κ=0.97) and low agreement for nitrates (κ=0.31) [66]. Similar findings were reported by Tisnado et al. who reported κ=0.80 for β- blockers and κ=0.20 for nitrates in their study [26]. Researchers have noted that this variation in agreement across drug classes is more likely due to the difference in the frequency and duration of medication use. For instance, β- blockers are used regularly, while nitrates are used only when needed [26, 66].

Studies have reported agreement for self-reported use of other prescription medications too. Agreement for gastrointestinal medication use was reported as moderate in various studies (κ=0.50 to κ=.067) [36, 68, 70]. Van de Brandt et al. reported 48% agreement for gastrointestinal drugs in their study [29].

High agreement has been documented for medications used for musculoskeletal disorders [50, 73]. In a cohort of patients with a diagnosis of rheumatoid arthritis (RA), Solomon et al. obtained overall moderate agreement for self-report of glucocorticoid use as compared to low agreement for disease-modifying antirheumatic drugs (DMARDs) use, for the reason that the former are comparatively more commonly used [50]. Specifically, they found high agreement for RA medications being used currently (κ=0.96 for methotrexate and κ=0.92 for hydroxychloroquine), but relatively lower agreement for RA medications used in the past (κ=0.13 for methotrexate and κ=0.35 for hydroxychloroquine) [50]. The authors reasoned that lower agreement for past RA medication use could be because the patients could not recall specific drug use [50]. In another study, Curtis et al. also reported high agreement for osteoporosis medication use in a cohort of glucocorticoid users, ranging from κ=0.80 (for alendronate), to κ=0.64 (for calcitonin) [73].

Fair-to-substantial agreement has been reported for self-report of antidepressant medication use (κ=0.42 to κ=0.77) [16, 26, 65, 68, 69]. Researchers have noted that agreement for antidepressants is usually lower because of the discomfort or reluctance of patients to report such medication use [16, 29, 70]. These findings illustrate the issue of social-desirability bias in reporting the use of medications for psychological conditions [16].

Many studies have also reported moderate-to-high agreement for HRT use (κ=0.51 to κ=0.92) [26, 68, 70-72]. One study however, reported lower agreement

(κ=0.21) for self-reported HRT use, by comparing self-report against data from prescription database [16]. The authors ascribed this low agreement to the rationale that the some pharmacies from where medicines were purchased were not covered by the particular prescription database used in the study [16].

For prescription NSAID use, fair-to-moderate agreement has been reported in the literature (κ=0.30 to κ=0.63) [26, 68, 70, 74]. Tisnado et al. reasoned that lower agreement for NSAIDs use may be because for their short-term use [26]. West et al. further analyzed the influence of recall interval and found that agreement for NSAID use decreased with an increase in the recall interval [28]. They also found that the recall interval also influences the agreement for recall of specific drug name [28]. They reported that the name of NSAID medication used 2-3 years before the interview was recalled more often than those used 7-11 years prior to the interview [28].

For self-reported opioid analgesic use, poor agreement has been reported in the literature. Caskie et al. reported κ=0.15 for opioid analgesic use and attributed this to the low prevalence of opioid analgesic users in their study population [68]. Tisnado et al. and Nielsen et al. also reported low agreement for opioid analgesic use (κ=0.40 and κ=0.49 respectively) [26, 70]. Low agreement in these studies has been ascribed to the fact that opioid analgesics are often used intermittently or short-term and, therefore, are not easily recalled by patients [26, 70].

Substantial agreement for antidiabetic medications has been previously reported. Studies have reported agreement for insulin in the range of κ=0.60 to κ=0.78 [26, 70]. Agreement for self-report of any oral hypoglycemic use has been previously reported as

κ=0.75 [70]. Tisnado et al. reported agreement for various classes of oral hypoglycemics as κ=0.60 for thiazolidinediones (TZDs), sulfonylurea and biguanides respectively [26]. However, Glintborg et al. reported a much higher agreement (κ=0.93) for sulfonylureas in their study [39].

Summary:

In Summary, the findings of above-mentioned studies indicate that prescription medications used for chronic conditions are recalled more accurately than medications used for acute conditions or on a short-term basis. Also, agreement for prescription medication use differs by the therapeutic class of medication in question.

2.1.3 Predictors that influence the accuracy of recall

Studies that assessed patient characteristics with regard to accuracy of self-report reported that age, limited memory, fading cognitive ability, educational status, income, health status, level of physical activity, and duration or severity of disease can affect the accuracy of recall [15, 47, 75]. Mukerji et al. and Okura et al. have also reported that the presence of other comorbid conditions can influence the level of agreement [15, 47]. Okura et al. and Merkin et al. noted that agreement for self-report of presence of a comorbid condition is higher among females as compared to males [15, 47].

Most notable predictors for recall of prescription medication use reported in the literature include duration of medication use and recall interval [28, 29, 68, 70, 76]. West et al. noted that recall interval was a significant predictor of NSAID drug name recall.

Shorter recall interval was associated with higher recall of NSAID name and dose, compared to a longer recall interval [28]. Nielsen et al. and Van den Brandt et al. also reported that medications used for a longer period are better recalled [29, 70]. The Van den Brandt study reported better recall of drugs used for longer duration (63.6% agreement for drugs used for 24 months or longer; 65.9% agreement for cardiovascular drugs used for 24 months or longer) as compared to drugs used for a relatively short period of time, e.g., 59.0% for drugs used for 6-11.9 months and 48% agreement for alimentary tract disorder medications [29]. Caskie et al. and Kelly et al. also reported lower agreement for medications taken for less severe conditions as compared to those taken for more serious conditions [68, 76].

2.2 Assessment of agreement in pregnant women

This section is divided into four parts: 1) accuracy of report for risky behavior among pregnant women; 2) accuracy of report of pregnancy history and related events; 3) accuracy of report for prescription medication use among pregnant women; 4) predictors that influence accuracy of report among pregnant women.

2.2.1 Accuracy of report for risky behavior among pregnant women

Studies conducted among pregnant women to assess the accuracy of self-reported information for risky behaviors, have mostly focused on smoking status and alcohol consumption during pregnancy. Majority of the studies used serum/urine cotinine levels

as the ‘gold-standard’ to check the accuracy of self-report for nicotine use and medical records for alcohol.

One of the main concerns conveyed by authors is significant inaccurate reporting of information regarding risky behaviors by pregnant women. Ford et al. in their study found that 19.2% pregnant women identified themselves as active smokers during pregnancy, whereas results of the serum cotinine tests showed this proportion to be 31.3% [33]. In contrast, they found self-reported information given by pregnant women regarding their smoking status before pregnancy to be accurate [33]. Further, Webb et al. found that 73% of women in their study reported as not being an active smoker during pregnancy. However, this could not be corroborated with their serum urine cotinine level, which was more than the predetermined cut off (80 ng/ml) to identify active smokers [32]. Furthermore, Britton et al. estimated that 34.7% of pregnant women in their study inaccurately reported as being nonsmokers, while their urine cotinine levels indicated otherwise [77]. These findings truly point towards the issue of social-desirability bias among pregnant women while reporting information about smoking status.

Of particular concern are the findings of the study by Rice et al. in which they observed a good agreement for smoking status (κ=0.80), whereas poor agreement (κ=0.17) was found for alcohol use during pregnancy [78]. The authors based this low agreement for alcohol use to the fact that alcohol use was not regularly recorded in the medical records [78]. Similarly, Hessol et al. also found lower agreement for alcohol consumption in a study conducted in a cohort of pregnant Latina women [79]. They

estimated the kappa value (κ) for self-report of alcohol use during pregnancy to be 0.37 using medical records as the 'gold-standard' [79].

However, findings of a few studies show that sometimes pregnant women do provide accurate information about smoking and alcohol consumption. This was shown in the study by Klebanoff et al. where the authors compared self-reported information about smoking given by pregnant women, with their serum cotinine levels and found that majority (87%) of pregnant women identified as active smokers during pregnancy, provided accurate information (κ=0.83) [42]. In another study conducted in a cohort of pregnant women participating in the NICHD Trial of Calcium for Pre-eclampsia Prevention (CPEP), Klebanoff et al. found substantial agreement (κ=0.72) between self-reported information and serum and urine cotinine levels [35]. Yawn et al. also a reported high agreement (κ=0.85, 93% agreement) for self-reported smoking during pregnancy upon retrospective comparison with medical records [31].

In another study Fox et al. assessed reliability of self-reported information about smoking status and alcohol consumption in a cohort of pregnant women participating in a randomized clinical trial of smoking cessation [80]. They compared self-reported information about smoking status and alcohol consumption prior to pregnancy that was collected first at 15±3.8 weeks of gestation, and later in the eighth month of gestation. It was compared with the thiocyanate levels from their saliva samples collected at the time of the first interview [80]. The results of the study showed that the agreement for smoking status was identical in both the intervention (κ=0.61) and the control group (κ=0.56) of the trial [80]. Similar results were also obtained for alcohol consumption

(κ=0.52 for intervention group and κ=0.55 for control group respectively) [80]. Authors however suggested that awareness among the study participants regarding verification of their self-reported information against their saliva samples might have influenced the level of agreement in a favorable way [80].

Researchers have also used other sources of data to ascertain agreement for self-reported smoking and alcohol use. Ernhart et al. and Jacobson et al. used information collected at two different points in time, i.e., during pregnancy and post-partum [34]. Ernhart et al. reported that most women under-reported the information about alcohol consumption during pregnancy, i.e., they reported lower level of drinking when asked during pregnancy compared to ascertainment 5 years after their pregnancy [34]. Jacobson et al. in their study conducted in a cohort of black pregnant women to estimate accuracy of reporting alcohol consumption, both during pregnancy and postpartum, have also indicated lower levels of agreement for ascertainment during pregnancy [81]. The authors suggested that this could be due the stigma associated with drinking during pregnancy [81].

2.2.2 Accuracy of report for pregnancy history and related events

Studies have been conducted to explore the extent of agreement for maternal recall of pregnancy related events, primarily in pregnant women considered to be in 'high risk' group for adverse perinatal outcomes. High agreement has been reported for recall of previous live births [16, 82-84], previous pregnancies and miscarriages [16, 82-85]. Low agreement for self-report of complications during pregnancy has also been documented [82, 86, 87].

2.2.3 Accuracy of report for prescription medication use among pregnant women

Very few studies have been conducted to assess the agreement for prescription medication use among pregnant women. Overall, poor-to-moderate agreement for prescription medication use has been reported [30, 40, 85, 88].

Olesen et al. conducted a study in Denmark to assess agreement for prescription medication use in a cohort of 2,041 pregnant women [40]. They used self-reported data from interviews conducted first at 6-12 gestational weeks and later at 12-15 gestational weeks, and compared these data with the information from a county prescription database [40]. They reported higher recall rates for prescription drugs dispensed 30 days prior to the interview (50%, 95% CI:43;46), compared with the drugs dispensed 120 days prior to the interview (43%, 95% CI:40;46) [40]. They also found that women had a higher recall for prescription medications used for treatment of chronic conditions (100% agreement for insulin, thyroid drugs, antiepileptic, and cardiovascular drug use; 80% agreement for antidepressant medication use, and 76% agreement for asthma medication use) as compared to those used on a short-term basis (40% agreement for NSAID use; 35% agreement for anti-infectives use, 47% for antacid use, and 59% agreement for antihistamine use) [40].

Bryant et al. compared self-reported information about the types of medications used (prescription medication, OTC medication and vitamins/supplements) and specific short-term illnesses/conditions during pregnancy with medical records, in a cohort of 202 still-pregnant and post-partum women [30]. They obtained moderate agreement for

prescription medication use during pregnancy (κ=0.48), but lower agreement for OTC medication and vitamin/supplements use (κ=0.02 and κ=0.07 respectively) [30]. Further, they noted that agreement was somewhat lower for episodic illnesses that occurred during pregnancy, e.g., κ= 0.07 for flu and upper respiratory infection and κ= 0.11 for nausea/vomiting [30]. These low levels of agreement obtained in this study were attributed to the reason that pregnant women do not view short-term illnesses significant enough to be reported to the physicians [30].

In another study de Jong et al. assessed agreement for specific classes of prescription medications used in a cohort of 246 post-partum women. They compared information about medication use that was collected at the time of the participants' prenatal visits and compared it with information collected retrospectively seven years later [88]. Overall, the authors found moderate agreement (55%) for medications used throughout pregnancy [88]. Specifically, they found that the agreement was highest for medications used during labor and delivery (77%) [88]. Moreover, de Jong et al. also reported higher agreement for medications for which a list of name of the drugs was provided. This was also corroborated by Mitchell et al. who noted that recall is influenced by the nature of question asked [89].

2.2.4 Predictors that influence accuracy of self- report among pregnant women

Various predictors that affect agreement for reporting risky behavior have been documented, however very few predictors that influence accuracy of reporting prescription medication use during pregnancy have been reported. Klebanoff et al.

studied the effect of race and ethnicity on recall and reported that African-American women provide more accurate information about smoking behavior than white women (κ=0.90 and 0.80, respectively) [42]. Hessol et al. reported that Spanish speaking Latina pregnant women in their study cohort were more likely to accurately report alcohol use during pregnancy, compared with English speaking Latinas [79]. Jacobson et al. and Ernhart et al. noted that history of alcohol abuse and maternal depression might influence agreement for alcohol and drug use during pregnancy [34, 81]. Britton et al. have reported that multigravidity, multiparty and number of smokers in a household can influence the accuracy of reporting smoking status during pregnancy [77]. For prescription medication use, recall interval and type of illness (chronic/acute) has been reported to influence the rate of recall [30, 40].

2.3 Summary of the literature review

The existing literature both in general population and pregnant women provides an insight into various factors that can potentially influence the self-reported information provided by patients. Even though heterogeneity of study samples makes comparisons across studies difficult, frequently or commonly used medications are recalled easily [28, 38, 50]. So are medications used for longer duration (chronically) [28, 70, 73], as compared to medications that are "used when needed" or short-term/intermittently [66, 68, 70]. In patients with chronic diseases, recall accuracy is reported to be influenced by the number of medications they are taking and also the presence of other underlying conditions that they might have [29, 38]. In addition, perceived 'chronicity' and

‘severity’ of a condition also makes its recall easier, as conditions that patients identify as having more severe effects on their lives and daily activities, are recalled accurately [12, 29, 38, 54].

Studies have demonstrated that recall of past medication use is also affected by the nature of the question asked. Typically, agreement is higher when question about the use of specific medication is asked for self-report than collecting this information through an open-ended question [15, 26, 38, 48, 50, 70, 89, 90].

Studies have reported inconsistent results in assessing accuracy of risky behavior during pregnancy. But there is substantial evidence that pregnant women are uncomfortable in revealing information regarding smoking and alcohol use during pregnancy, because of social stigma and disgrace attached to this behavior, as they are considered inappropriate during pregnancy. They however report such behavior accurately when asked retrospectively. Specifically in pregnant women, there is evidence that recall accuracy regarding medication use might be influenced by the recall interval.

*** Refer to Appendix 1 for detailed description of the studies mentioned above in a tabular form.**

CHAPTER-3 STUDY METHODS

This chapter discusses the research design, study methodology, hypothesis testing, and statistical analyses in detail.

3.1 Study design

The study was a cross-sectional analysis to assess the validity of self-reported information provided by pregnant women on prescription drugs that they take during pregnancy. This was achieved by comparing their self-reported information (obtained through a standardized questionnaire) about prescription medication use relative to the information present in their medical records (the 'gold standard').

Since for the purposes of this validation study subjects were not followed up, a *cross-sectional study design* was utilized. For the purposes of this study, the information regarding prescription medication use derived from the patients' electronic medical records, was assumed to be the 'gold standard', as it is the most comprehensive source of patients' medical information, including information regarding prescription medications and various inpatient and outpatient records. Hence, comparison of responses of pregnant women regarding their prescription medication use with the information in their electronic medical records was the most appropriate way to assess the validity of the information they give for their prescription medication use. Alternative approaches and limitations of the 'gold-standard' used, i.e., medical records are presented in the Discussion section (Chapter-5).

3.2 Self-reported data- the 'SMART' study

The study was conducted by utilizing data from The Safety of Medications During Pregnancy and Women's Perception of Teratogenicity (SMART) study, an ongoing prospective cohort study being conducted at the University of New Mexico (UNM), Albuquerque. The SMART study was initiated to ascertain the safety of the most common medications (prescription, OTC, herbal products, and dietary supplements). The study also aims to determine perception of teratogenic risk (perceived hazard) of the medications taken by the study participants, and also ascertain barriers of patient-provider communication regarding medication use during pregnancy. The study is approved by UNM Human Research Review Committee (HRRC).

Study participants for the SMART study were recruited from the UNM Main Hospital (UNMH) and its five satellite clinics throughout the city of Albuquerque, New Mexico. These satellite clinics are affiliated with the UNM Department of Obstetrics and Gynecology (OB/GYN). Patients at these clinics seek preventive medicine, family planning, prenatal and postnatal care and treatment for chronic diseases, e.g., diabetes mellitus, hypertension.

Pregnant women attending these prenatal clinics were contacted by a healthcare provider and their interest for participation in the study was sought. Pregnant women who were $\geq$ 18 years old, and had no prenatal diagnosis indicating an abnormal pregnancy were included in the study. Women were recruited at any time of gestation and were willing to be interviewed in either English or Spanish. All the study participants gave a

written informed consent. As of 11/05/09, of the 494 women approached for participation, 406 agreed to participate, resulting in a participation rate of 82.4%. Lack of interest and time constraints were the most common reasons cited to choose not to participate.

Self-reported data

All the participants participated in a semi-structured interview (Appendix 2) of about 20-25 minutes duration, in English or Spanish, depending on their preference of language. These interviews were administered by a trained bilingual interviewer, who at the time was a Ph.D. candidate in Anthropology at UNM. The participants also granted permission to access their medical records and permission to contact them with a follow-up phone interview, if needed.

In the interview, the participants were asked to report all the prescription, over-the-counter (OTC), herbal products and dietary supplements that they took since their last menstrual period (LMP). For ascertaining prescription medication use, the participants were asked, " Have you taken any medications prescribed by your doctor or any other healthcare provider since your last menstrual period, even if you stopped taking them once you knew you were pregnant?" They were also asked to indicate their perception of teratogenicity about these medications on a 5-point Likert-type scale, where 1 indicated 'not likely to cause harm' and 5 indicated 'very likely to cause harm'. In addition, participants were also asked about their knowledge and attitude towards medication use during pregnancy, and the sources of information that they refer to. Additionally, information on general demographic and lifestyle characteristics, medical and

reproductive history, and any pregnancy complication/s was also collected. Collected data were entered into an SPSS® database.

3.3 Electronic medical record review- The 'Gold-Standard'

The SMART study database also contains information about subjects' prescription medication use, which is systematically abstracted from their electronic medical profiles (PowerChart®, Cerner Corporation). The PowerChart® is an electronic health record management system developed by the Cerner Corporation® and universally used by all services of UNMH. The patients are identified by their Medical Record Number (MRN) which is provided by the University of New Mexico Hospital (UNMH) to abstract information regarding prescription medication use from PowerChart®.

The medications identified from the PowerChart® records were classified into the drug classes of interest and then recorded into the database. One respondent could have multiple medication use, thus information about multiple medications was recorded under respective fields in the database. For example, if a patient reports a use of glyburide, it will be recorded in the class "sulfonylurea" for diabetes medication. If the same patient reported the use of codeine, it was recorded in the "moderate/low agonist" class for opioid analgesics.

In the medication list of PowerChart®, medications were listed as either "ordered" or "documented". While "ordered" medications are those that are prescribed by a physician at UNMH, "documented" medications are the ones that are reported by the

patient during the medication reconciliation process, and not necessarily consist of medications prescribed by a physician at UNMH. For the purposes of this study, information about "ordered" prescription medications was abstracted from the PowerChart®, given the higher accuracy of such entries.

3.4 Study population and sample selection

The study population consisted of the first 311 pregnant women enrolled in the SMART study. For the selection of final study sample the following eligibility criteria were used:

3.4.1 Eligibility criteria:

Inclusion criteria:

1. Pregnant women enrolled in the first year of the SMART study.
2. Women who had at least one prescription for medication for either diabetes or opioid analgesics from a provider at UNMH or its affiliated clinics (List of specific classes is provided in Table 3 and 4).

Exclusion criteria

1. Prescription of opioid analgesics for chronic illnesses or chronic pain (requiring treatment/medication for three months or more) [91, 92].

Rationale for inclusion/exclusion criteria:

The first 311 pregnant women enrolled in the SMART study were chosen because these women had complete data derived from both the SMART study questionnaire and PowerChart®. Women having at least one prescription for diabetes or opioid analgesics

were chosen because we were interested in checking the validity of self-report among pregnant women in chronic conditions vs. acute conditions.

3.4.2 Definition of chronic medication use

The National Center for Health Statistics (NCHS) defines a chronic condition as the one that cannot be prevented by vaccines or cured by medicines [91]. Chronic conditions continue for long durations, usually lasting for more than three months and require long term medication use [91, 92]. For the purposes of this study, the above presented definition was used to classify the medication use as chronic or acute. *Chronic medication use* was defined as the medication use lasting for 3 months or more, whereas *acute medication use* was defined as use for short-term or episodic conditions. Diabetes was the most common chronic condition in this population. Therefore, antidiabetic medications were chosen to represent chronic medication use.

Analgesics can be given for acute/episodic (e.g., fever, inflammation, pain resulting from headache or backache) or for chronic conditions (e.g., rheumatoid arthritis, chronic migraine). Even though NSAIDs were the most widely used class of analgesics in this population, they were not included in this study as they are mostly available as over-the-counter (OTC) products. Therefore, opioid analgesics given for acute or short-term use (< 3 months) were chosen as they can be obtained only through prescription.

After identification of a drug class from the self-reported data and medical records, prescription medication use was categorized into one of the classes considered in

this study, i.e., antidiabetic medication and opioid analgesics. Detailed exposure information for these prescription medications are as follows:

1. **Diabetes medication (Table 3):** Study participants taking prescription medication from at least one of the four major classes of the diabetic medication [8, 93].
2. **Opioid analgesics (Table 4):** Study participants taking prescription medication from at least one of the four major classes of opioid analgesics [8, 93].

Table 3: Drug classes for antidiabetic medication

Drug class	FDA risk classification	Route of administration
Insulin and its analogs *(injectables)*		Parenteral (subcutaneous or intravenous)
a. Lispro insulin solution	B	
b. Insulin aspart	B	
c. Insulin glulisine	C	
d. Insulin glargine solution	C	
e. Insulin detemir solution	C	
Sulfonylureas		Oral
a. Tolbutamide	C	
b. Glipizide	C	
c. Glyburide	C	
d. Glimepiride	C	
Biguanides		Oral
a. Metformin	B	
Thiazolidinediones (TZDs)		Oral
a. Pioglitazone	C	
b. Rosiglitazone	C	

Table 4: Drug classes for opioid analgesics

Drug class	FDA risk classification	Route of administration
Strong agonists		Oral/parenteral (intravenous)
Alfentanil	C	
Fentanyl	C	
Heroin	B	
Meperidine	B	
Morphine	C	
Oxycodone	B	
Remifentanil	C	
Sufentanil	C	
Hydromorphone	B	
Oxymorphone	B	
Moderate/Low agonists *(available in combination with acetaminophen)*		Oral
Codiene	C	
Propoxyphene	C	
Hydrocodone	C	
Mixed agonists-antagonists and partial agonists		Oral/parenteral (intravenous)
Butorphanol	C	
Nalbuphine	B	
Pentazocine	C	
Other		Oral
Tramadol	C	

The initial study population consisted of the first 311 women enrolled in the SMART study. Records (questionnaire data) of one patient did not have information regarding prescription medication use, as this patient had to leave midway during the interview because she started experiencing labor pains. Therefore, this patient was excluded from the study. 310 patients with complete records (both self-report and

medical records) constituted the study sample. Electronic medical records of all the 310 patients were reviewed for ascertaining documentation of prescription medication use for either antidiabetic medication or opioid analgesic medication. After the review process, a total of 92 patients were found to have a "recorded" medication use either, in self-report or in the electronic medical record. Women who did not answer an interview question (missing data) were excluded from the corresponding analyses in which that information was needed.

3.5 Study variables

The study variables, obtained from the database (including derived/dummy variables) and used in the statistical analyses are described below:

Table 5: Study variables

VARIABLE NAME	DESCRIPTION	CATEGORIES
Demographic variables		
Age	Maternal age at the time of interview	18-23 years 24-29 years 30 years and above
Educational level	Maternal educational level at the time of interview	Less than high school High school/GED College and above
Marital status	Maternal marital status at the time of interview	Single, never married Married/Living with partner Separated/Divorced/Widowed
Insurance coverage	Type of insurance coverage at the time of interview	No insurance Have any insurance
Ethnicity	Maternal ethnicity	White, non-Hispanic White, Hispanic Others
Place of birth	Maternal place of birth (United States or outside United States)	United States Outside United States
Language	Primary language	English Spanish Other

VARIABLE NAME	DESCRIPTION	CATEGORIES
Lifestyle characteristics		
Smoking status	Current maternal smoking status	Never smoked Past smoker Current smoker
Binge drinking	Alcohol consumption of 4 or more drinks since LMP on a single occasion	Yes No
Knowledge and attitude factors		
Medication use during pregnancy	Knowledge and attitude towards medication use during pregnancy	Stop taking *all* the medications upon recognition of pregnancy Continue with *necessary medications* Continue taking *all* the medications as needed
Birth defects caused by medications	Knowledge about ability of medications taken during pregnancy to cause birth defects	Never Sometimes Often Very Often Always
Alcohol consumption during pregnancy	Knowledge and attitude towards alcohol consumption during pregnancy	Should abstain OK to consume some alcohol OK to drink wine/beer, but not hard drinks
Sought consultation about safety of medication	Knowledge and attitude towards medications currently prescribed (during pregnancy)	Yes No
Medical/Reproductive Factors		
Chronic conditions	Presence of chronic conditions	None At least one
Gestational age	Gestational age at the time of interview	Less than 20 weeks More than 20 weeks
History of previous adverse perinatal outcome	Adverse perinatal outcomes before the current pregnancy, e.g., miscarriage, ectopic pregnancy, termination, stillbirth	Yes No
Gravidity	Number of times the respondent has been pregnant	Primigravid Multigravid
Parity	Number of times the respondent has given live births	Nulliparous Parity more than one

3.6 Data analysis

3.6.1 Outcome measure: agreement for self-reported prescription medication use

The primary outcome to be measured in this study was the extent of agreement between the responses provided by pregnant women regarding their prescription medications for chronic or acute use with the information present in their electronic medical records. For this purpose the agreement was defined as the presence of *concordant* responses about prescription medication use in self-report and medical records. This is illustrated in Table 6.

Table 6: Representation of agreement/disagreement

Information about prescription medication from SELF-REPORT	Information about prescription medication from MEDICAL RECORD	Outcome
Yes	Yes	Agreement
Yes	No	Disagreement
No	Yes	Disagreement

In this study, agreement for chronic medication use was represented by agreement for antidiabetic medication use, and that for acute or short-term use, was represented by agreement for opioid analgesic use. Agreement was measured by estimating a simple kappa statistic (κ), prevalence-adjusted and bias-adjusted kappa (PABAK) and observed proportion of agreement (P_o) values.

For assessing agreement for self-reported prescription medication use of antidiabetic medications, agreement was first assessed for the specific drug classes that it

comprises of, as identified in Table 3. For this, simple kappa, PABAK and the observed proportion of agreement values were calculated for each class. Thereafter, a commutative value (mean value) was calculated for agreement of antidiabetic medication use. Similarly, agreements for all the drug classes for opioid analgesics identified in Table 4 were calculated, followed by calculation of a mean value of agreement measures. This mean value represented agreement for overall self-reported use of opioid analgesic medication

Table 7 illustrates an example, whereby it is shown how the information regarding self-reported prescription medication use and the information derived from the medical records was recorded in the database. In this hypothetical example, patient#001 reports the use of glyburide (antidiabetic medication) in self-report. This is recorded in the class 'sulfonylurea-self-report' for diabetic medication use. If this same information is corroborated from the PowerChart® medication profile of the patient, then it is reported in 'sulfonylurea-medical record' of the patient, and according to Table 6 the outcome is agreement. In the case of patient#002, information from self-report regarding use of opioid analgesic use is oxycodone. It is classified under the class 'moderate opioid-self-report', but in the medical records, this information is not present, hence according to Table 6, it is counted as a disagreement. Since a patient can have multiple drug use, patient #003 represents how such a case might look like, if a patient is reporting the use of all the three medications.

Table 7: Example showing representation of self-reported data and medical record data in the database

Pat id	Su SR	Su MR	In SR	In MR	Mo SR	Mo MR	O Su	O In	O Mo
001	1	1	0	0	0	0	1	-	-
002	0	0	0	0	1	0	-	-	0
003	1	0	1	1	1	1	0	1	1

Where, pat id= patient's id

1= Yes

0 = No

SR = Self-Report

MR = Medical Records

Su = Sulfonylurea

In = Insulin

Mo = Moderate

O = Outcome of interest, i.e., agreement on medication use; (1= agreement, 0= disagreement)

Sample characteristics were summarized by descriptive statistics. Frequencies for categorical data and means or medians for continuous data were presented.

3.6.2 Kappa statistic (κ)

In this study, agreement was estimated by comparing kappa statistic calculated for

prescription medication use for chronic (represented by diabetes medication use) and acute (represented by opioid analgesics) conditions. Kappa statistic is a common measure of validity and reliability of categorical data which takes into account the agreement that can occur due to chance. It can have values from -1 to 1. Several classifications have been proposed for interpreting kappa values and this study would utilize the classification proposed by Landis and Koch [94]. According to this classification, if kappa value lies between 1.0 and 0.8 then it is regarded as almost perfect agreement [94]. If the kappa value lies between 0.8 and 0.6 then it is considered a substantial agreement, whereas if the value lies between 0.6 and 0.4 then it is considered to be a moderate agreement. Kappa value between 0.4 and 0.2 is considered as fair agreement, value between 0.2 and 0 is considered slight, and kappa value below zero is poor agreement [94]. This is shown below in Table 8.

Table 8: Landis and Koch's classification for interpretation of a kappa value

Range of kappa value (κ)	**Interpretation**
1.0 – 0.8	Almost perfect
0.8 – 0.6	Substantial
0.6 – 0.4	Moderate
0.4 – 0.2	Fair
0.2 - 0	Slight
0 - -1	Poor

Usually, for calculation of kappa statistic, a 2 X 2 table is used as shown below in Figure 1. Cells 'a' and 'd' represent the number of cases where there is agreement between both the gold-standard and the observer for the presence and absence of the

outcome of interest. Cells 'b' and 'c' represent the number of cases where the gold standard and the observer do not agree on the presence of the outcome. Calculation of kappa statistic takes into consideration the difference between the agreement actually present (proportion of observed agreement, P_o) and the agreement that is present by chance alone (proportion of expected agreement, P_e). It is calculated as follows:

	Gold Standard (Medical Records)		
Observer (Self-report)	**Yes**	**No**	**Total**
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	**N**= a+b+c+d

Figure 1: A 2X2 table for calculation of kappa statistic

Kappa statistic (κ) = $\frac{\text{observed agreement } (P_o) - \text{expected agreement } (P_e)}{1 - \text{expected agreement}}$

Where, observed agreement, $P_o = (a+d) / N$

expected agreement, $P_e = [(a+c)(a+b) + (b+d)(c+d)] / N^2$

3.6.3 Prevalence-adjusted and bias-adjusted kappa (PABAK)

Even though kappa statistic is widely used to measure agreement between two raters, it suffers from certain limitations, most notable of them being its dependence on prevalence [95, 96]. If the horizontal and vertical marginal cells of a 2 X 2 table are relatively unbalanced, i.e., the prevalence of the desired outcome is either very low or high, the value of kappa statistic maybe misleading. For instance, it may signify a low

level of agreement, even if the proportion of observed agreement is high. Feinstein and Cicchetti have described this phenomena as the kappa paradox [95]. To correct this, Byrt et al. proposed the use of prevalence-adjusted and bias-adjusted kappa (PABAK) [96] . The effect of prevalence can be assessed by calculating the prevalence index (PI). It is calculated by the following formula:

$$PI = \frac{|\ a-d\ |}{N}$$

If the prevalence index is high, the expected agreement is also high and the corresponding kappa value is lower [96, 97]. Bias index (BI) represents the extent to which propensity of two raters to classify the occurrence of the outcome in yes or no categories differs [96]. It is calculated as:

$$BI = \frac{|\ b-c\ |}{N}$$

PABAK adjusts the kappa statistic for the influence of high and low prevalence by substituting the actual values of cells ‘a’ and ‘d’, by their average values. Similarly, for adjusting the bias introduced due to different observers, values for cells ‘b’ and ‘c’ are replaced by their average values. The 2X2 contingency table then looks like this:

		Gold Standard (Medical Records)	
Observer (Self-report)		**Yes**	**No**
Yes		p	q
No		q	p

Where, $p = (a+d) / 2$ and,

$q = (b+c) / 2$

Using these values to calculate PABAK, we get [96]

$$PABAK = 2P_o - 1$$

An alternate version of this equation, using the prevalence index and bias index values gives [96],

$$PABAK = \kappa (1- PI^2 + BI^2) + PI^2 - BI^2$$

Hoehler has criticized the use of PABAK citing that adjusting for prevalence and bias effects might result in overestimation of kappa [97]. Our choice to use PABAK was driven by the inherent limitations in the kappa statistic, which PABAK takes care of. While there is no consensus as to what measure of agreement should be reported, researchers have advocated that it is more insightful to present more than one measure of agreement [97]. In our study, there was a considerable prevalence effect due to no data in the cell 'd' (medical record- "No", self-report- "No").

3.6.4 Sensitivity and specificity

In addition to measures of agreement, sensitivity and specificity of the maternal self-report in comparison with information in medical records (*the 'gold standard'*) were also calculated as measures of validity. Sensitivity and specificity are common measures used to assess the indicative ability of any test [98]. While sensitivity is the ability of any test to correctly identify those who have the outcome of interest, specificity is the ability

of any test to correctly identify those who do not have the outcome of interest. Sensitivity is calculated as the number of true positives divided by a combination of the number of true positives and the number of false positives. Specificity is calculated as the number of true negatives divided by a combination of the number of true negatives and the number of false negatives). This is illustrated in the following figure:

	Gold Standard (Medical Records)		
Observer (Self-report)	**Yes**	**No**	**Total**
Yes	a TP	b FP	a+b
No	c FN	d TN	c+d
Total	a+c	b+d	

Where, TP = True positives

FP = False positives

TN = True negatives

FN = False negatives

Sensitivity = TP/ (TP + FN) = a/ (a+c)

Specificity = TN/ (FP + TN) = d/ (b+d)

Figure 2: Sensitivity and specificity

3.7 Hypothesis testing

Statistical analyses conducted to test the study hypotheses are presented below:

Research hypothesis 1: Agreement between self-reported information and information in

the medical records will be greater for medications prescribed for chronic medical conditions (represented by diabetes medication), compared with prescription medications given for acute medical conditions (represented by opioid analgesic) in the SMART study cohort.

Unit of analysis: Major therapeutic class of chronic and acute medication as identified for this study.

Kappa statistic, prevalence-adjusted and bias-adjusted kappa (PABAK) and observed proportion of agreement (P_o) values were calculated for each class of antidiabetic and opioid analgesic medications as identified in Table 3 and 4.Value of kappa statistics and PABAK for each class was then pooled by taking their overall mean (commutative) kappa value for prescription medication for diabetic medication. This is illustrated in the formula below:

$$\text{Mean kappa for diabetic medication use} = \frac{\text{Sum of kappa values from all classes}}{\text{Number of drug classes}}$$

Similarly, mean kappa was computed for opioid analgesic medication as:

$$\text{Mean kappa for opioid analgesics} = \frac{\text{Sum of kappa values from all classes}}{\text{Number of drug classes}}$$

The mean kappa value was used to compare the difference in agreement between chronic and acute medication use. Sensitivity and specificity of self-reported information by the patients about prescription medication use were also calculated for antidiabetic and opioid analgesic drug use by using information from the medical records as the 'gold-standard'.

Research hypothesis 2: To examine the predictors of inaccurate reporting in this UNM-based study cohort.

Unit of analysis: Study subjects with inaccurate information about prescription medication use.

A multivariate logistic regression analysis was performed to identify the significant predictors of inaccurate reporting in this study. The outcome of interest for this analysis was inaccurate report of prescription medication use. For this, inaccurate reporters for each medication class were first identified. These inaccurate reporters consisted of those patients who had discordant information regarding prescription medication use (Table 6).

Therefore, according to the example in Table 7, pat#002 is an inaccurate reporter for insulin use, and pat#003 is an inaccurate reporter for sulfonylurea use. Thus in other words, a patient was considered to be an inaccurate reporter if the information about medication use was present in the medical records but not in the self-report, or vice versa.

In the case multiple medication use, patients were classified as inaccurate reporters if they inaccurately reported medication use for even one class of medication, even if they had concordant information for the rest of the classes of medication. This method has previously been used in studies assessing recall of medical conditions, where the researchers employed the above mentioned method to identify inaccurate reporters for reporting diagnosis of chronic conditions [15, 48, 60, 65]. To identify inaccurate reporters among multiple medication users, an alternative method of classification for accurate/inaccurate reporting of medication use would have been to have multiple

observations per person, i.e., one observation for each medication use (for e.g., if a patient is on five different classes of medications, that patient would be present in the database five times, once for each class), and then assess the accuracy of recall. However, in this method the observations would no longer be independent and it would violate the assumption of the logistic regression analysis of independent observations. This would have resulted in erroneous results. Therefore, inaccurate reporters were classified according to the method used in previous studies. This method does not violate the assumptions of logistic regression and the observations are independent of each other. Dependent variable was dichotomous (inaccurate report; yes=1/no=0).

Since no prior studies have been conducted among pregnant to assess the predictors of inaccurate recall regarding medication use, demographic factors were chosen based on prior studies done in the general population. Medical/reproductive factors were included based on prior studies conducted among pregnant women that assessed predictors of medication use. Health literacy is defined as the ability of patients to read, understand and act on medical instructions. Level of health literacy of a patient can influence the extent of agreement for medication use. Knowledge and attitude factors, along with educational status reflect the extent of health literacy among the study participants and were used as proxy measures of health literacy.

Inaccurate reporters from all the medication classes were then pooled together and constituted the study sample for the analyses for this study hypothesis. The following representation shows how this looked like in the database.

Pat id	Inaccurate reporter
001	0
002	1
003	1

Univariate analyses using the Chi-square/ Fisher's exact test and t-test methods (as appropriate) were performed to identify any significant difference between the groups (inaccurate reporters and accurate reporters). Potential predictors were tested for significance at $p<0.20$ in univariate analyses [99]. Thereafter, multivariate regression analysis was conducted to estimate if the selected covariates predict inaccurate reporting for prescription medication use. Covariates included in the final multiple regression model included: insurance status (no insurance as reference category), place of birth (place of birth outside United States as reference category), at least one episode of binge drinking ($\geq$ 4 drinks/occasion) around LMP (no episode of binge drinking as reference category), presence of chronic conditions (no chronic conditions as reference category), and gestational age at the time of interview (gestational age $\leq$ 20 weeks as reference category), and number of unique prescription medications.

Significance was tested at $p<0.05$. All the analyses for hypotheses testing were conducted by using SAS software (version 9.1.3).

CHAPTER 4 - RESULTS

This chapter presents the results of the study. A description of the study population is presented, followed by the results for the measures of agreement. Finally, the results of the multiple logistic regression analysis with respect to important predictors of inaccurate report are presented.

4.1 Description of the study sample:

Results from the descriptive analyses are presented in Tables 9-11.

4.1.1 Demographics and lifestyle characteristics

Demographic and lifestyle information is presented in Table 9. The mean age of the study participants was 29.2 ± 6.1 years and half of the patients (50%) were more than 30 years old. More than three-fourths of the participants had a high school or higher degree (77.2%), with about 31.5% reporting having a vocational or college degree. In contrast, about 22.8% indicated that they had educational experience of less than high school. About 40.2% of the participants were either married or living with a partner and almost an equal proportion (39.1%) of them were separated, divorced or widowed. About a third (33.7%) of the study participants did not have any insurance coverage.

More than half (54.3%) of the study participants spoke English as their primary language, whereas 43.5% indicated Spanish was their primary language. A majority (59.8%) of the participants were born in the United States and were of Hispanic ethnicity (67.4%).

About 65.2% of the study participants reported that they had never smoked cigarettes or used tobacco. About a third of the women (32.6%) indicated that they were past smokers (quit smoking upon recognition of pregnancy or before that). Only two women (2.1%) reported at the time of the interview that they were still active smokers (current smokers). Almost a third of the women (29.4%) reported that they had engaged in 'binge drinking' (consumption of 4 or more drinks at a single occasion), at least once in the periconceptional period.

4.1.2 Knowledge and attitude towards medication use during pregnancy

Table 10 shows the results about knowledge and attitude factors regarding medication use during pregnancy in this study cohort. Majority of the participants (86.9%) had sought consultation about the safety of medications they took during pregnancy. Majority of the participants (70.3%) reported that upon recognition of pregnancy, consumption of 'only necessary' medications should be continued, while about a quarter (26.4%) of them indicated that all the medication should be stopped upon recognition of pregnancy. Among study participants, majority had accurate knowledge about how often medication use can during pregnancy cause birth defects ('sometimes': 70.4%), while about a quarter (28.3%) of the participants had exaggerated perception of teratogenicity and indicated that medications used during pregnancy can cause birth defects 'often', 'very often' or 'always'. A majority of the women (94.6%) in the study sample also indicated that consumption of alcohol should be stopped upon recognition of pregnancy.

Table 9: Demographic and lifestyle characteristics of the study population (N=92)

Variable	N (%)*
DEMOGRAPHIC FACTORS	
Age	
18-23 years	21 (22.8)
24-29 years	25 (27.2)
30 years and above	46 (50.0)
Educational level	
Less than high school	21 (22.8)
High school/GED	32 (34.8)
Some College/Vocational school	29 (31.5)
College degree	07 (7.6)
Masters, doctorate or professional degree	03 (3.3)
Marital status	
Single, never married	19 (20.7)
Married/Living with partner	37 (40.2)
Separated/Divorced/Widowed	36 (39.1)
Insurance status	
No insurance	31 (33.7)
Have any insurance	61 (66.3)
Ethnicity	
White, non-Hispanic	15 (16.3)
White, Hispanic	62 (67.4)
Other	15 (16.3)
Place of birth	
United States	55 (59.8)
Outside United States	37 (40.2)
Primary language	
English	50 (54.3)
Spanish	40 (43.5)
Other	02 (2.1)
LIFESTYLE FACTORS	
Smoking status	
Never smoked	60 (65.2)
Current smoker	02 (2.2)
Past smoker	30 (32.6)
• Quit before pregnancy recognition	19 (20.6)
• Quit after pregnancy recognition	11(12.0)
At least one episode of binge drinking (≥ 4 drinks /occasion) around last menstrual period (LMP)	
Yes	27 (29.4)
No	65 (70.7)

* Sample size might vary due to missing data

** Decimal points have been rounded off

Table 10: Knowledge and attitude about medication use and alcohol consumption during pregnancy (n=92)

Variable	N (%)*
Attitude towards medication use during pregnancy	
Stop taking all the medications upon recognition of pregnancy	24 (26.4)
Continue with necessary medications	64 (70.3)
Continue taking all the medications as needed	03 (3.3)
How often can medications cause birth defects	
Never	01 (1.2)
Sometimes	57 (70.4)
Often	12 (14.8)
Very often	04 (4.9)
Always	07 (8.6)
Attitude towards alcohol consumption during pregnancy	
Should abstain	87 (94.6)
OK to consume some alcohol	04 (5.4)
OK to drink wine/beer, but not hard drinks	0
Sought consultation about safety of medication	
Yes	80 (86.9)
No	12 (13.0)

* Sample size might vary due to missing data
** Decimal points have been rounded off.

4.1.3 Medical and reproductive history

The results of medical and reproductive history are presented in Table 11. Nearly three-fourths of the women (73.9%) reported at least one chronic condition at the time of interview. More than half (54.3%) of the women reported a diagnosis of preexisting or gestational diabetes. This can be explained by the sampling procedure utilized for the study, i.e., selection of participants having a prescription for either antidiabetic or opioid analgesic medication which contributed towards a high prevalence of chronic conditions, especially diabetes, in this cohort. Other most commonly reported chronic conditions reported in this study sample were migraine headaches (17.4%), asthma/allergies (10.9%) and depression (9.8%).

The mean gestational age of the study participants at recruitment was 32.5 ± 5.9 weeks, with 93.5% of subjects recruited after 20 weeks gestation. A third (29.7%) of the study participants had a history of adverse perinatal outcomes. Prior adverse perinatal outcomes included: miscarriage (32.6%), stillbirth (1.08%), termination (9.78%) and, ectopic pregnancy (2.17%). About 71.4% of participants reported that they had given birth to a live child before (multiparous), while for 17.6% of women, this was the first time they were pregnant (primigravida).

Table 11: Medical and reproductive history of the study sample (n=92)

Variable	**Mean± S.D.**
Gestational age at the time of interview (weeks)	32.5 ± 5.9
Variable	**N (%)***
Presence of chronic conditions	
None	24 (26.1)
At least one	68 (73.9)
History of adverse perinatal outcomes	
Yes	27 (29.7)
No	64 (70.3)
Gravidity	
Primigravid	16 (17.6)
Multigravid	75 (82.4)
Parity	
Nulliparous	26 (28.6)
Parity more than one	65 (71.4)
Gestational age at the time of interview	
Less than or equal to 20 weeks	06 (6.5)
More than 20 weeks	86 (93.5)

* Sample size might vary due to missing data

4.2 Measures of agreement and validity

4.2.1 Prevalence of major therapeutic classes use

Prevalence of medication use was estimated for both data sources used, i.e., self-report from the SMART study questionnaire and the medication use data abstracted from the patients' electronic medical records (EMRs). No recorded use was found for 'mixed agonist-antagonist/partial agonist class' and 'other opioid analgesic class' for opioid analgesics, therefore they were not included in any analyses to assess the measures of agreement. Average duration of opioid analgesic use among study participants was 5.3 days. None of the study participants were found to be chronic users of opioid analgesics as per the definition of chronic users for the purposes of this study. Thiazolidinedione (TZD) class for antidiabetic medication had only one recorded use from the medical records; therefore it was also excluded from the main analyses for estimating the measure of agreement. The results are presented in Table 12.

Among antidiabetic medication use, sulfonylurea had the highest recorded use (29.4%), followed by insulin and its analogues (23.9%) and biguanides (22.8%). For opioid analgesic use, the highest recorded use was found for strong agonists (42.4%), followed by moderate/low agonist use (22.8%).

Among antidiabetic medication, almost equal proportion of use was found for insulin (21.7%; for both self-report and medical record) and biguanides (21.7% use documented in self-report vs. 20.7% documented use in the medical records). For sulfonylureas, considerable difference in the documented use in the two data sources was found (22.8% use in self-report vs. 29.7% use documented in the medical records).

Similar results were obtained for use of strong agonist class of opioid analgesics, i.e., data from the medical records indicate that about 39.1% of the sample were prescribed strong opioid analgesics, while only 23.9% reported its use by self-report.

Table 12: Prevalence of Medication Use by Reporting Source (N=92)

Drug class	Self-report N (%)	Medical record N (%)	Total prevalence N (%)
ANTIDIABETIC MEDICATION			
Insulin and its analogues	20 (21.7)	20 (21.7)	**22 (23.9)**
Sulfonylureas	21 (22.8)	27 (29.4)	**27 (29.4)**
Biguanides	20 (21.7)	19 (20.7)	**21 (22.8)**
Thiazolidinediones (TZDs)	0	1 (1.1)	**01 (1.1)**
OPIOID ANALGESIC MEDICATION			
Strong agonists	22 (23.9)	36 (39.1)	**39 (42.4)**
Moderate/Low agonists	10 (10.9)	13 (14.1)	**21 (22.8)**
Mixed agonist-antagonists and partial agonists	0	0	**0**
Other	0	0	**0**

4.2.2 Measures of agreement/concordance

For measures of agreement, simple kappa and 95% confidence interval, prevalence-adjusted-bias-adjusted kappa (PABAK), and proportion of observed agreement (P_o) were estimated. Information regarding medication use from the medical records was used as the gold-standard. Agreement between self-report and medical records varied by drug class, as shown in Table 13. Overall, almost perfect agreement was found for antidiabetic medication use (mean kappa=0.87 and mean PABAK=0.91) in

contrast to poor-to-moderate agreement for opioid analgesics (mean kappa=0.29 and mean PABAK=0.57).

Among antidiabetic medications, agreement was highest for biguanide use (κ=0.90, CI=0.79; 1.00, PABAK=0.9, P_o=0.97), followed by insulin use (κ=0.87, CI=0.75; 0.99, PABAK=0.93, P_o=0.95) and sulfonylurea use (κ=0.83, CI=0.70; 0.96, PABAK=0.87, P_o=0.93). For opioid analgesic use, while PABAK and proportion of observed agreement (P_o) were similar for the two classes included in the analysis (strong agonist: PABAK=0.65, P_o=0.78; moderate/low agonist: PABAK score=0.59, P_o=0.79), simple kappa varied significantly (κ for strong agonists=0.51, 95% CI=0.33; 0.69, κ for moderate/low agonists=0.06, 95% CI= - 0.17; 0.29).

Table 13: Measures of concordance/agreement* (N=92)

Drug Class	Simple kappa (95% CI)	p-value	PABAK**‡	Proportion of observed agreement (P_O)
ANTIDIABETIC MEDICATION				
Insulin and its analogues	0.87 (0.75;0.99)	<0.001	0.93	0.95
Sulfonylureas	0.83 (0.70;0.96)	<0.001	0.87	0.93
Biguanides	0.90 (0.79;1.00)	<0.001	0.93	0.97
OPIOID ANALGESIC MEDICATION				
Strong agonists	0.51 (0.33;0.69)	<0.001	0.56	0.78 (78.2%)
Moderate/Low agonists	0.06 (-0.17;0.29)	0.29	0.59	0.79 (79.3%)

* Medical record information is the "gold-standard".
‡ 95% CI for PABAK not calculated.
*** 'Thiazolidinedione', 'Mixed agonist-antagonists /partial agonists' and, 'other opioid' classes were excluded from analyses because of empty cells i.e., no reported use either in self-report or medical records.
**** Decimal points have been rounded off.

Table 14: Sensitivity and specificity measures using medical records as the 'gold-

Drug class	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)
ANTIDIABETIC MEDICATION *		
Insulin	0.90 (0.68-0.98)	0.97 (0.90-0.99)
Sulfonylurea	1.00 (0.83-1.00)	0.92 (0.82-0.97)
Biguanides	0.90(0.68-0.98)	0.99 (0.93-0.99)
OPIOID ANALGESIC MEDICATION **		
Strong opioid agonist	0.86 (0.65-0.97)	0.76 (0.64-0.85)
Moderate opioid agonist	0.20 (0.02-0.55)	0.87 (0.77-0.93)

***** Sensitivity and specificity not calculated for Thiazolidinedione' because of small sample size.
** Sensitivity and specificity not calculated for 'Mixed agonist-antagonists /partial agonists' and, 'Other opioid' because of empty cells i.e., no reported use either in self-report or medical records.
*** Decimal points have been rounded off.

4.2.3 Sensitivity and specificity measures

Sensitivity and specificity scores with 95% confidence interval were estimated. The results are shown in Table 14. Overall, good sensitivity and specificity measures were obtained for antidiabetic medication use. While sensitivity and specificity were similar for insulin and biguanides use, sulfonylurea use had higher sensitivity but lower specificity. For opioid analgesic medication use, 86.3% sensitivity was obtained while specificity was 75.7%. Use of moderate opioid agonists was associated with a very low sensitivity of 20%, whereas specificity was 86.5%.

4.3 Predictors of inaccurate reporting of prescription medication use

Significant covariates identified in the univariate analysis included: maternal age , insurance status, place of birth, at least one episode of binge drinking since LMP, presence of chronic conditions, gestational age at the time of interview, and the number of unique prescription medications per patient (prescription burden). After controlling for all the other covariates in the model, episode of binge drinking and presence of chronic conditions was found to be significantly associated with inaccurate reporting of prescription medication use. Specifically, women who reported at least one episode of binge drinking since their LMP were 3.4 times more likely (95% CI: 1.13; 10.29) to inaccurately report prescription medication use during pregnancy.

Table 15: Results of univariate tests (N=92)

Variable	Accurate reporters (N=45) N (%)	Inaccurate reporters (N=47) N (%)	p-value
Age			<0.0001*
18-23 years	08 (17.8)	13 (27.7)	
24-29 years	12 (26.7)	13 (27.7)	
30 years and above	25 (55.6)	21 (44.7)	
Educational level			0.97
Less than high school	11 (24.4)	10 (21.3)	
High school/GED	16 (35.6)	16 (34.0)	
College and above	18 (40.0)	21 (44.7)	
Marital status			0.31
Single, never married	08 (17.8)	11 (23.4)	
Married/Living with partner	20 (44.4)	17 (36.2)	
Separated/Divorced/Widowed	17 (37.8)	19 (40.4)	
Insurance status			0.02*
No insurance	18 (40.0)	13 (27.7)	
Have any insurance	27 (60.0)	34 (72.3)	
Ethnicity			0.57
White, non-Hispanic	06 (13.3)	09 (19.1)	
White, Hispanic	31 (68.9)	31 (66.0)	

Variable	Accurate reporters (N=45) N (%)	Inaccurate reporters (N=47) N (%)	p-value
Other	08 (17.8)	07 (15.0)	
Place of birth			0.005*
United States	22 (48.9)	33 (70.2)	
Outside United States	23 (51.1)	14 (29.8)	
Primary language			0.28
English	20 (44.4)	30 (63.8)	
Spanish	25 (55.6)	15 (31.9)	
Other	0	02 (4.3)	
Smoking status			0.49
Never smoked	32 (71.1)	28 (59.6)	
Past smoker	13 (28.9)	17 (36.2)	
Current smoker	0	02 (4.3)	
At least one episode of binge drinking (≥ 4 drinks /occasion) around LMP			0.005*
Yes	07 (15.6)	20 (42.6)	
No	38 (84.4)	27 (57.4)	
Attitude towards medication use during pregnancy			0.32
Stop taking all the medications upon recognition of pregnancy	12 (26.7)	12 (25.5)	
Continue with necessary medications	29 (64.4)	35 (74.5)	
Continue taking all the medications as needed	03 (6.7)	0	
How often can medications cause birth defects			0.28
Never	01 (2.2)	0	
Sometimes	31 (68.9)	26 (55.3)	
Often	04 (8.9)	08 (17.0)	
Very often	03 (6.7)	01 (2.1)	
Always	02 (4.4)	05 (10.6)	
Attitude towards alcohol consumption during pregnancy			0.62
Should abstain	43 (95.6)	44 (93.6)	
OK to consume some alcohol	02 (4.4)	02 (4.3)	
Sought consultation about safety of medication			1.00
Yes	39 (86.7)	41 (87.2)	
No	06 (13.3)	06 (12.8)	
Presence of chronic conditions			0.03*
None	07 (15.6)	17 (36.2)	
At least one	30 (66.7)	38 (80.9)	
History of adverse perinatal outcomes			1.00
Yes	13 (28.9)	14 (29.8)	

Variable	Accurate reporters (N=45) N (%)	Inaccurate reporters (N=47) N (%)	p-value
No	32 (71.1)	32 (68.1)	
Gravidity			0.95
Primigravid	07 (15.6)	09 (19.1)	
Multigravid	38 (84.4)	37 (78.7)	
Parity			0.86
Nulliparous	13 (28.9)	13 (27.7)	
Parity more than one	32 (71.1)	33 (70.2)	
Gestational age			<0.0001*
Less than or equal to 20 weeks	02 (4.4)	03 (6.4)	
More than 20 weeks	43 (95.6)	44 (93.6)	
Number of unique prescription medications (prescription burden)	**Mean (SD)** 4.7 (3.5)	**Mean (SD)** 6.1 (5.4)	<0.0001*

*= significant at α= 0.2.

** Decimal points have been rounded off.

Table 16: Predictors of inaccurate reporting of prescription medication use

Maternal characteristic/predictor	Adjusted Odds ratio	95% CI	p-value
Age			0.75
18-23 years	1.0*	--	
24-29 years	1.32	0.33; 5.28	
30 years and above	1.76	0.45; 6.98	
Insurance status			0.27
No insurance	3.5	0.32; 37.03	
Have any insurance	1.0*	--	
Place of birth			0.13
United States	5.80	0.57; 58.79	
Outside United States	1.0*	--	
At least one episode of binge drinking (≥ 4 drinks /occasion) around LMP			**0.03****
Yes	**3.40**	**1.13; 10.29**	
No	1.0*	--	
Presence of chronic conditions			0.07
None	1.0*	--	
At least one	0.28	0.08; 1.00	
Gestational age			0.52
Less than or equal to 20 weeks	1.0*	--	
More than 20 weeks	1.83	0.29; 11.42	
Number of unique prescription medications (prescription burden)	1.03	0.94; 1.138	0.49

Note:

1. * = Reference category
2. ** = significant at $\alpha = 0.05$
3. All odds ratio are adjusted for all variables in the table.
4. Non significant covariates as identified in the bivariate analysis were not included in the final model.
5. Decimal points have been rounded off.

CHAPTER 5 DISCUSSION

This chapter presents a discussion about the descriptive results of the study followed by discussion on the results of the main analyses. Limitations of the study are then discussed, followed by implications and recommendations for future research.

5.1 Discussion of descriptive findings

The study was conducted in a predominantly Hispanic population (67.4%). This is particularly noteworthy because none of the earlier studies to assess agreement for prescription medication use among pregnant women have been conducted in a primarily Hispanic population. The sample was recruited from UNM Main Hospital and five community clinics that are situated throughout the city of Albuquerque, New Mexico, which allowed for capturing a socio-economically diverse sample and reduce selection bias. The interview for obtaining self-reported information about prescription medication use was conducted both in English and Spanish, therefore omitting the potential confounding due to language.

One finding worth noting in this study sample was that majority of the women had educational experience of high school and above (77.2%), of which 7.6% had a college degree. While high educational status has been reported as a predictor of higher recall for prescription medication use, this was not the case in this study sample. About three-fourths of the sample (66.3%) had some kind of insurance coverage. Lack of insurance coverage has been reported as predictor of low agreement for prescription medication use. In our study, however, we were unable to find a significant association between the two.

Women in this study sample, in general had more accurate knowledge regarding alcohol consumption and medication use during pregnancy. About 70 % women reported that upon recognition of pregnancy, only necessary medication should be taken after consultation with a medical practitioner. This is in accordance with the current guidelines on prescription medication use during pregnancy [2, 9, 25]. About 87% of the women in the sample had sought consultation about the medication they were taking during pregnancy. This can be explained in part, by our sample selection. Since, we were interested in assessing accuracy of prescription medication use; women in our sample had some type of prescription medication, and must have sought consultation about the medications that were prescribed to them. Majority of the women (94.6%) reported that alcohol consumption should be stopped during pregnancy. This is also in line with the guidelines by U.S. Surgeon General's advisory on alcohol use during pregnancy [100].

5.2 Measures of agreement

The primary objective of this study was to assess whether agreement for prescription medication use is influenced by the type of medication use (chronic vs. acute). The findings of the study indicate that agreement for antidiabetic medications, i.e., chronic use, was much higher as compared to that for opioid analgesics, i.e., acute use. While almost perfect agreement was obtained for all the classes of antidiabetic medications, poor-to-moderate agreement was obtained for opioid analgesic use.

Even though previous research has shown that accuracy of recall for prescription medication use differs according to the therapeutic category under investigation [29, 50,

68, 70], various studies conducted in the general population have consistently reported a higher level of agreement for prescription medications used for a longer duration or 'chronically' [26, 29, 68, 75, 76]. Consistent with the findings of these studies, we obtained high agreement for prescription medications used chronically to control diabetes ($\kappa = 0.87$; PABAK = 0.91). Agreement for specific drug classes of antidiabetic medications has also been previously reported in the literature among non-pregnant women. Moderate-to-substantial agreement for insulin and its analogues (κ =0.6 to κ =0.78) [26, 70] and oral hypoglycemics (κ =0.75) [70] has been reported. Specifically, κ =0.60 has been reported for biguanides and TZDs [26], whereas for sulfonylureas moderate-to-excellent agreement has been reported (κ =0.60 [26] and κ =0.93 [39]. Our study demonstrates even higher agreement for these specific drug classes (insulin: κ =0.87, PABAK=0.93; sulfonylurea: κ =0.83, PABAK=0.87; biguanides: κ =0.90, PABAK=0.93). Our kappa value for sulfonylurea medications was however lower than the one obtained by Glintborg et al [39].

For self-reported opioid analgesic medication use among non-pregnant populations, low-to-moderate agreement has been reported in the literature in the range of κ =0.15-0.49 [26, 68, 70]. In accord with these previously reported values, we also found low agreement for opioid analgesic use in our sample (κ =0.29; PABAK=0.57). Agreements for specific classes of opioid analgesics have not been reported in the literature; therefore we were unable to make any comparisons.

Among pregnant women, only a few studies have reported agreement for prescription medication use [30, 40, 85, 88]. These earlier studies have used different

methodologies to ascertain agreement for prescription medication use. Bryant et al. used medical records as the 'gold-standard' to assess accuracy of self-report of prescription medications used in a cohort of pregnant women [30]. While their choice of 'gold-standard' was similar to ours, their method of ascertaining information about prescription medication use was different in that they asked for report of 'any' prescription medication use. The results of this study therefore, only reveal information on the agreement of use of prescription medication among pregnant women.

In another study conducted by de Jong et al., the researchers assessed the agreement by estimating sensitivity and specificity measures for specific therapeutic categories (including antibiotics, antacids, anesthetics, hypnotics/sedatives, tocolytics, and oxytocics), that did not include our therapeutic categories of interest (i.e., antidiabetics and opioid analgesics) [88]. In addition, their data sources and 'gold-standards' were also different from what we have utilized in our study. They used information on medication use collected at the time of prenatal visits by pregnant women, and compared it with data collected seven years later [88]. Their choice of 'gold-standard' has a serious flaw, as recall diminishes with increase in the recall interval [16, 28, 70, 71]. This might have caused considerable recall bias in the study.

However, the study conducted by Olesen et al. improved upon these methodological flaws by utilizing prescription database as the 'gold-standard', which provided with the information about the medications that were actually dispensed [40]. Even though this information does not suggest that the patient actually consumed the medication, it at least provides with an indicator that the medications were dispensed and

may have been used by the patient. Olesen et al. assessed agreement for the most commonly used prescription medications in their cohort [40]. This was the only study that assessed agreement for specific classes of medication. They found perfect agreement (100%) for insulin use [40]. This was the only medication class that we could compare to in previous reports. These results are also in accordance with our posited hypothesis that medications used for longer duration (i.e., antidiabetic medications) are recalled accurately.

However, our results are more robust than Olesen et al, as we calculated kappa statistic that takes into account agreement that can occur due to chance, while Olesen et al estimated agreement percent agreement. Further, we also estimated PABAK that adjusts for dependence of the kappa statistic on prevalence of the outcome if interest.

5.3 Predictors of inaccurate reporting

Our results indicate that women who had at least one episode of binge drinking were more likely to inaccurately report the use of prescription medications, independent of other factors like age, insurance status, place of birth, presence of chronic conditions gestational age, and number of unique prescription medications. About 30% of our study participants reported at least one binge drinking episode ($\geq$ 4 drinks per occasion) before pregnancy recognition a month around LMP. This finding holds a great significance given that heavy drinking during pregnancy can result in fetal alcohol spectrum disorder (FASD). Alcohol consumption during pregnancy itself is considered a risky behavior. If

binge drinking further leads to inaccurate reporting of prescription medication use, then it calls for increased surveillance over such women regarding the information they provide for prescription medication use. Previous research indicates that high alcohol intake is associated with lower agreement regarding prescription medication use among women [101]. Merlo et al in their study conducted in a cohort of Swedish women found that women who engaged in high alcohol consumption were 1.47 times more likely (95% CI; 1.09:1.97) to inaccurately report their prescription medication use, as compared with women who did not engage in high alcohol consumption. While this study could not demonstrate that alcohol consumption can lead to inaccurate reporting of prescription medication use as the women included in this study were slightly older (45-73 years), it does demonstrates that alcohol intake can influence agreement regarding medication use among women in general.

Since during medication reconciliation, it is not possible to ascertain that the recorded medications are actually consumed, medical records of women engaging in alcohol consumption or binge drinking, especially the information that they provide regarding their prescription medication use should be supplemented with other sources of information, e.g., pharmacy records.

Age, educational status, insurance status, and number of medications used are the other most commonly reported predictors for accuracy of recall for prescription medication use [29, 71, 75, 102]. In this study, however we were unable to obtain significant association with these commonly reported predictors with respect to accuracy of recall of prescription medication use.

5.4 Limitations

The findings of this study should be interpreted in light of a few limitations associated with the study design. These limitations are presented below:

1. There might have been variability in patient recall based on the interval between the timing of the interview and the timing of prescription medication use. The mean gestational age at enrollment was 32.5 weeks ± 6.0. Most of the women in the study sample were interviewed in the late second or early third trimesters. Recall interval has been reported to be a significant predictor of agreement for prescription medication use in the general population [28]. Olesen et al. and Bryant et al. have also reported that the recall interval can influence recall among pregnant women [30, 40]. However, in this study we did not assess the association between the recall interval and accuracy of report.
2. Recall can also be influenced by the route of medication used by the patient, as the patient is more likely to remember those medications for which they have to follow specific instructions, over those that do not have any such instructions for administration, such as, parenteral vs. oral medications. This has been previously reported in the literature as one of the factors [68, 73].
3. Due to the small sample size (92 patients), power of the study was limited and did not allow for comparisons by specific medications. Therefore, for the purposes of this study, agreement was ascertained by the major class of medications instead of specific medications. Comparison by specific drugs would have been more accurate but was not logistically feasible due to the small sample size.

4. In this study, we were assessing validity for the most common chronic and acute medication use that are used during pregnancy. Therefore, the results of this study cannot be generalized to classes of medications other than those utilized in this study.
5. The study results have limited generalizability since it was conducted at one location only (UNMH, Albuquerque); thus, the study sample might not be representative of the entire U.S. population.

6. ***Use of electronic medical records***

i. In this study, medical records were used as the 'gold-standard' for assessing accuracy of self-reported information about prescription medication use. Researchers argue that no source of information has complete and accurate information regarding medication use [65, 103]. While self-reported information is often affected by various information biases, medical records have also been criticized for not being a complete source of information for medication use [26, 36, 44]. However, they play an important role in clinical decision making process [48]. Nowadays, with the increasing use of electronic medical record (EMR) system, most of the challenges and limitations earlier associated with medical records, e.g., delayed and erratic recording, illegible handwriting of physicians, non-documentation of medications taken from an outside provider, and very limited information about lifestyle and behavioral characteristics, are taken care of [104, 105]. Apart from streamlining and improving the documentation of patient related data, data in EMRs can also be

used in clinical research, and safety and quality assurance studies [104, 105]. UNM Hospital has been using the PowerChart® system by Cerner Corporation for managing the medical information and prescription medication lists. In this study, therefore, medical records were used as the 'gold-standard' as they were the most comprehensive and updated source of information regarding the medication use.

ii. One limitation associated with this study was absence of information about prescription medications prescribed by a provider other than those at UNMH. In PowerChart, only information about medication prescribed by a UNMH provider is present. It is possible that patients enrolled in the study could have seen other providers outside of UNMH. Information about prescription medications prescribed by those outside providers was not present in PowerChart, and was available only through patient self-report. 43 patients (47%) were found to be taking prescription medication from an outside provider. Further, recent immigrants among this study sample might have brought prescription medication from other countries. This can influence the false-positive rate (i.e., patients who do not have a documentation of prescription medications in the medical records, but reported the use of medication in self-report). An increase in the number of false-positives can reduce the *specificity* of the study. In addition, it is probable that in the medication reconciliation process, information regarding some prescription medications might not have been taken out from the medication profiles of the

patient even if they were discontinued by the patient. This could influence the false-negative rate (i.e., information about medications present in medical records, but not reported by the patients). An increase in the number of false-negatives can reduce the *sensitivity* of the study. Both of these cases can result in a decrease in the level of agreement. However, it was not found to be the case in this study. As can be seen in Table 12, the proportion of self-reported medication users medication users identified from medical records was similar. Further, this was also corroborated from the results of sensitivity and specificity analysis (Table 14), which shows almost perfect sensitivity and specificity for these two measures.

iii. Finally, for the purposes of this study only the information regarding documentation of prescription of antidiabetic and opioid analgesics was abstracted from the electronic medical records (PowerChart®) of the patients. This documentation of medication is done by the physician attending to the patient. This information is not suggestive of information on compliance to those medications by the patients. In the self-report, the patients only report if they have taken prescription medications that they were prescribed by their doctor since their last menstrual period. We did not capture adherence, i.e., checked if the medications were taken as prescribed. As obtaining this information is not possible from the PowerChart®, this information might go uncaptured. This might influence the agreement of medication use, as those patients who report having received a prescription of a certain medication

might not have actually used that medication.

5.5 Implications for future research

Future studies should be designed to address the limitations mentioned earlier. Further studies are needed to explore agreement for other commonly used classes of medications. As suggested by the findings of this study, patient self-report cannot be relied upon for medications used for short-term or intermittently. Thus, additional sources should be used to validate the self-reported information, e.g., pharmacy databases. Prescription databases and pharmacy claims are reported to be more accurate sources to validate information about prescription medication use as compared to medical records [50, 73, 106], since they provide accurate information whether or not a prescribed medication was filled, but information about the actual exposure is still lacking. If the sample size permits, analyses should also be done for recall accuracy of specific medications rather than classes.

Results of this study also provide with some insight on the patient characteristics associated with inaccurate report of prescription medication use. Results of this study suggest that information regarding prescription medication use provide by pregnant women without any chronic conditions and engaging in atleast one episode of binge drinking may be unreliable. Using these information physicians can develop effective methods of identifying pregnant women who have these characteristics and are more likely to inaccurately report their prescription medication use.

5.6 Recommendations

The reliability of the self-reported information on medication use has importance in the clinical practice and research where accurate information regarding drug use/exposure is required. Self-report is the most common method of obtaining information about medication use. Relying on patient self-report alone can lead to incorrect estimation of medication exposure, especially for medications taken for short-term.

With respect to research in the field of safety of medication in pregnancy, often self-report is the only method of obtaining information regarding drug exposure. Lower agreement for medications used short-term suggests that self-report for such medication is unreliable and can lead to misclassification of exposure to medications and erroneous assessment of teratogenicity. Therefore, other sources like prescription databases to validate the self-reported medication use are needed. Use of such data sources would also help in capturing non-compliance, especially in women who are more likely to inaccurately report their medication use.

Another key area that requires attention is specificity of the questions about medication use. Previous research demonstrated that agreement depends on the nature of question asked. If women are asked about the use of a particular medication or given a list of specific medications, they are more likely to accurately recall medication use as

compared to situations when they are presented with open-ended questions about medication use.

In this study, women responded to an open-ended question about prescription medication use, since they were recruited from multiple specialty clinics and a potential list of medications would be too long. Recall of medication use could have been better, especially for opioid analgesic medications, if the question was more specific and specific brand of generic names of medications were provided.

5.7 Conclusion

The findings of this study suggest poor accuracy of self-report with respect to prescription medications used as short courses or intermittently during pregnancy. Accuracy of report for prescription medications taken chronically was higher. Pregnant who engaged in binge drinking were more likely to inaccurately report use of medications. Therefore, in clinical studies assessing safety of such medications in pregnancy, self-reported information needs to be supplemented by other sources.

APPENDIX 1: DESCRIPTION OF STUDIES IN A TABULAR FORM

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
STUDIES CONDUCTED IN GENERAL POPULATION					
Haapanen et al. (1997) [13]	Middle aged and elderly cohort (n=596)	Medical records	Hypertension, myocardial infarction, angina pectoris, coronary heart disease, claudication, cerebral stroke, diabetes, hip/knee arthrosis, lower back disorder	High agreement was noted for conditions like diabetes, hypertension, myocardial infarction (κ=0.78, 0.78, 0.77 respectively), whereas lower agreement was found for conditions like lower back pain (κ=0.42).	Had predefined classification criteria for identification of medical conditions in the patients' medical record for which agreement was to be assessed. Results showed higher agreement for chronic conditions and those conditions that have well defined diagnostic criteria.
Bush et al. (1989) [14]	Elderly cohort (n=120)	Medical records	Angina, cancer (any), cataracts, diabetes, fractures, hypertension, myocardial infarction, stroke	Good agreement was observed for most conditions. Highest agreement were noted for diabetes (κ=0.93) and stroke (κ=0.85).	No assumed gold-standard. The study reported an overall high percentage agreement, which might have been due to participation of the study participants in an earlier screening program.
Kriegsman et al. (1996) [12]	Elderly cohort (n=2,380)	Information from general practitioners ("Alloyed gold-standard")	Chronic non-specific lung disease, cardiac disease, peripheral atherosclerotic disease, cerebrovascular disease, diabetes, malignant neoplasm, osteoarthritis/rheumatoi	High agreement for diabetes (κ=0.85), moderate agreement for neoplasm and cardiac disease (κ=0.64 and κ=0.69 respectively) was observed, whereas low agreement for osteoarthritis/rheumatoid	Information from general practitioners was considered to be an "alloyed", not true gold-standard. The sample chosen for this study was relatively sicker than the parent study's sample, which might have

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
			d arthritis	arthritis (κ=0.31) and peripheral atherosclerotic disease (κ=0.38).	contributed to increase in awareness regarding the medical condition present in participants. Findings of the study indicate that accuracy for recall is higher for chronic conditions that have a well defined diagnostic criteria and are easily understood by patients, whereas conditions with diagnostic criteria that are not easily understood by patients, have lower recall rates. Also, in this cohort it was observed that patients had higher recall for those chronic conditions that they perceived to be life threati9nging to them over those conditions that are generally associated with old-age and are not considered life threatening (example, arthritis)
Okura et al. (2004) [15]	General population (Olmstead county residents)	Medical records	Heart failure, diabetes, myocardial infarction, hypertension, stroke	Good agreement was found for diabetes (κ=0.76), hypertension (κ=0.75), myocardial infarction (κ=0.80) and stroke	Good study design. Using randomly selected population based sample minimized selection bias. Authors were able to justify

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	(n=2,037)			(κ=0.71), whereas lower agreement was found for heart failure (κ=046).	the use of medical records as the 'gold-standard'. Findings of the study also illustrate that accuracy for recall is higher for chronic conditions that have a well defined diagnostic criteria and are easily understood by patients, whereas conditions with diagnostic criteria that are not easily understood by patients, have lower recall rates.
Paganini-Hill et al. (1982) [16]	Elderly women cohort (n=334)	Medical records, Pharmacy records	Gallbladder disease, hypertension, diabetes, benign breast disease, hysterectomy, oophorectomy	Good overall agreement was observed for most conditions. Highest agreement was observed for hysterectomy (κ=0.96) and lowest agreement was observed for benign breast disease (κ=0.63).	Good study design (case-control).
Smith et al. (2008) [17]	Military cohort (n=37, 798)	Medical records	38 medical conditions; ranging from hypertension, chronic bronchitis, kidney failure, to sinusitis, depression, asthma etc.	Typically good agreement was noted for chronic conditions, for example, 53.5% positive agreement was found for hypertension and only 1.4% positive agreement was found for migraine.	The study cohort was relatively young and physically active in this study. The results suggested that use of electronic medical records might be appropriate to estimate prevalence of conditions in a sample.

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
Simpson et al. (2004) [20]	Elderly women (n=1, 002)	Medical records, laboratory and physical test results, physician questionnaire	14 medical conditions; ranging from hip fracture, Parkinson's disease, to osteoporosis, arthritis	Overall good agreement wt was found for most conditions. Highest agreement was found for hip fracture (κ=0.96) and lowest agreement was found for arthritis (κ=0.24).	Good study design. Used multiple sources to compare self-repotted information. Used random sampling, thus minimizing bias.
Tisnado et al. (2006) [26]	Cohort of patients enrolled in managed care organization (n=1,270)	Medical records, patient self-report	Diagnosis of medical conditions (cancer, diabetes, asthma etc.), medication use, counseling and referrals, clinical services delivered.	Overall good level of agreement was observed for diagnosis of conditions and medication use but for clinical services, counseling and referrals lower level of agreement was observed.	Good study design; used bootstrapping to account for representation of patients in multiple items in questionnaire. Authors provide a good explanation of using two gold standards and potential sources of disagreement.
Miller et al. (2008) [63]	Cohort of patients from the Medicare Health Outcomes Survey (HOS) who were also eligible for Veterans Affairs (VA) care (n=7,953)	VA medical records	10 chronic conditions (diabetes, hypertension, chronic lung disease, arthritis, angina, congestive heart failure, myocardial infarction, stroke, and cancer).	Good level of agreement was observed for most conditions (75%). Highest agreement was observed for diabetes (κ=0.82) and lowest for chronic low back pain (κ=0.33).	Nice study design. The authors were able to explain the possible sources of disagreement.
Caskie et al.	Cohort of low-	PACE	Ten major therapeutic	Overall high levels	Use of 'brown bag' for

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
(2004) [68]	income elderly patients	(Pharmaceutical Assistance Contract for the Elderly) program's pharmacy records	classes of drugs; ranging from antihistamines, anti-infective, to cardiovascular, gastrointestinal and CNS agents. Specific selected classes of drugs within two major classes (cardiovascular and CNS agents) were also included.	agreement was noted for agreement for self-reported medication use when compared with pharmacy data. Highest agreement was observed for cardiovascular drugs (κ=0.83).	collecting self-reported data regarding medication use is noteworthy and eliminates chances of biases.
West SL et al. (1995) [28]	Cohort of patients enrolled in managed care organization (n= 560)	Pharmacy dispensation database	Use of nonsteroidal anti-inflammatory drugs and noncontraceptive estrogens.	Fair agreement in recalling any use of NSAID was noted (41%) of which 30 % could recall name of the medication, while only 15% recalled both name and dose of the medication.	Good study design. Use of pharmacy dispensation database as gold standard provided accurate estimation of drug use.
Van den Brandt et al. (1991) [29]	Cohort of patients enrolled in managed care organization (n= 270)	Pharmacy records data	Prescription medication use for general use.	Moderate agreement (61.2%) was observed. Highest agreement was noted for cardiovascular drugs (65.9%).	Nice study design; use of population level data provided enough power for analysis.
Ferrante et al. (2008) [44]	Cohort of patients enrolled in the SCOPE (Supporting	Medical record	Cancer related risk factors, screening tests and counseling	Highest agreement was noted for self-reported information self-reported information regarding diagnosis of cancer (96%),	The study provides good information on the various factors related to level of agreement in a cohort of cancer patients. Use of

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	Colorectal Cancer Outcomes through Participatory Enhancements)study			whereas agreement for self-reported smoking cessation counseling was low (41%)	medical records as the gold-standard was appropriate considering the outcomes of interest.
Westbrook et al. (1998) [36]	Cohort of patients with dyspepsia	Medical records	Self-reported diagnosis of dyspepsia and its management	Overall poor agreement was observed. Low agreement was noted for factors like duration of dyspepsia (κ=0.34) and number of medication taken (κ=0.28 and κ=0.31 for medications taken before and after endoscopy respectively).	Results of the study illustrate the issue of lack of knowledge about the nature of conditions and misunderstanding the diagnoses of their conditions.
Bergmann et al. (1998) [45]	Cohort of participants in the American Cancer Society's CPS-II study (n=65, 582)	Population based cancer registry data	Self-reported cancer diagnosis	Level of agreement for self-reported diagnosis of cancer varied by site. Highest sensitivity was noted for cancer for breast (0.91), prostate (0.90) and lung (0.90), whereas lowest agreement was noted for cancer of rectum (0.16).	Good study design. Use of population level study cohort and state level registry data ensured sufficient power for analysis.
Hall et al. (2004) [46]	Cohort of patients enrolled in three health	Medical records	Self-reported information on digital rectal examination, prostate specific	Overall good agreement was found for PSA test, sigmoidoscopy, and colonoscopy (κ=0.40-	--

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	maintenance organization (HMOs) (n=2,1 30)		antigen (PSA) test, fecal occult blood tests, sigmoidoscopy, and colonoscopy	0.80).	
Mukerji et al. (2007) [47]	Cohort of patients having head and neck cancer (n=458)	Medical records	Self-reported comorbidities due to cancers other than head and neck cancer	Good agreement was found in general. Highest agreement t was found for diabetes (κ=0.89) and lowest for arthritis (κ=0.11).	--
Goebeler et al. (2007) [49]	Cohort of elderly patients aged 90 years and older (n=209)	Medical records	Self-reported medical history	Overall moderate to fair level of agreement was observed for most conditions. Agreement was highest for Parkinson's disease (κ=0.74) and lowest for depression (κ=0.11).	Good study design. The study illustrates various predictors that can influence recall in elderly population for chronic conditions.
St. Sauver et al. (2005) [38]	Cohort of patients enrolled in Mayo Clinic	Medical records	Self-reported information on cardiovascular disease (CVD) and it's risk factors	Level of agreement varied for various CVD conditions. Highest agreement was found for high blood pressure (77.9%) and lowest agreement was found for medical problems related to peripheral arteries (31%).	The study illustrates that recall rates are higher for chronic diseases.
Solomon et al. (2007)	Cohort of patients	Medical records	Self-reported information on	Agreement varied by drug class and time of use, for	--

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
[50]	enrolled in the BRASS (Brigham Rheumatoid Arthritis Sequential Study) study		rheumatoid and medication use	example agreement for current use for methotrexate was (κ=0.96) whereas agreement for past use of methotrexate was (κ=0.13).	
Merkin et al. (2007) [51]	Cohort of end stage renal disease (ESRD) in CHOICE (Choices for Healthy Outcomes in Caring for End-stage renal disease)study (n=965)	Medical records, physician report	Self-reported information on eight comorbid conditions : congestive heart failure, myocardial infarction, cerebrovascular disease, angioplasty or coronary artery bypass graft surgery, hypertension, diabetes, chronic obstructive pulmonary and cancer	Highest agreement was recorded for diabetes (κ=0.93) and lowest agreement was recorded for chronic obstructive pulmonary disease (κ=0.20).	The study illustrates the concern about low agreement for those conditions that do not have a clear diagnostic criteria and high agreement for those conditions that have higher awareness among a particular cohort.
Boissonnault et al. (2005) [52]	Cohort of outpatient orthopedic surgery candidates (n=100)	Medical records, physician report	Self-reported information on patients' illness history, surgery and medication use.	Overall substantial agreement was found (κ=0.69). Highest agreement was found for skin cancer (κ=0.58) and lowest agreement was found for pneumonia (κ=0.27).	The authors contest the issue of questionnaire design and suggest that 'open-ended questions' might contribute towards higher agreement.
Iversen et al. (2007)	Cohort of age-sex stratified	Medical records	Self-reported information regarding	Level of agreement varied by conditions. Highest	--

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	sample (n=2, 318)		chronic respiratory disease and health services utilization	agreement was observed for pulmonary tuberculosis (κ=0.88), lowest agreement was observed for chronic bronchitis (κ=0.10).	
Skinner et al. (2005) [48]	Cohort of patients from the Veterans Health Study (VHS) (n=402)	Medical records	Self-reported information about five chronic medical conditions: diabetes, hypertension, obstructive lung disease, chronic low back pain, osteoarthritis of knee	High agreement was noted for diabetes (κ=0.84) and hypertension (κ=0.70).	Findings of this study indicate that recall rates are higher for chronic conditions that have well defined diagnostic criteria than those that have ambiguous diagnostic criteria. Also, the study illustrates the issue of misunderstanding of the diagnosis by patients that can possibly contribute towards lower agreement
Hessol et al. (2001) [103]	Cohort of women enrolled in WIHS (Women's Interagency HIV study) study (n=339)	County level AIDS surveillance data	Self-reported information about diagnosis of AIDS and AIDS related conditions.	Fair agreement was found for self-reported information about AIDS diagnosis and registry data (73%).	The level of agreement varied by specific conditions. This could have occurred because of inaccurate information about non-AIDS related conditions from the patients, which they might have thought are AIDS related, and are not present in registries.

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
Kropp et al. (2007) [71]	Cohort of patients enrolled in the MARIE (Mammacarcinoma Riskfactor Investigation) study (n=449)	Prescription records from physicians	Self-reported hormone therapy use	Good agreement was observed (88.2%) for self-reported hormone therapy compared with prescription records form physicians.	--
Wang et al. (2003) [106]	Cohort of patients enrolled in an HMO and VAMC and with a diagnosis of hypertension (n=200)	Pharmacy prescription data	Self-reported information about missing antihypertensive medication	Poor agreement was noted for self-reported compliance with antihypertensive therapy (κ=0.12).	--
Sandini et al. (2008) [102]	Cohort of patients enrolled in the Kuopio Osteoporosis Study (OSTPRE) (n=11,377)	National level prescription database	Self-reported information about hormone therapy use	Good agreement (97.6%) was observed for self-reported hormone therapy use when compared with national level prescription database.	Good study design. The specific questions asking for self-reported hormone therapy use were clear and comprehensive. The questions also asked for names of hormone therapy medications, which contributed to high specificity.
Haukka et al. (2007) [69]	Cohort of patients having a diagnosis of	National level prescription	Self-reported information about prescription	Good overall agreement was observed for most psychotropic drugs.	Good study design using of national level database. Quite contrary to general

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	schizophrenia (n=905)	database	psychotropic medication	Highest agreement was observed for lithium (κ=0.96) and antipsychotics (κ=0.87).	assumption that agreement for psychotropic medications is lower, the study had high agreement for these drugs. Participants were asked to bring their medication prescription for psychotropic medications, which might have resulted in high overall agreement rates in this population.
Nielsen et al. (2008) [70]	Cohort of patients from the Danish health survey for the year 2000 (n=16,688)	National level prescription records	Self-reported information regarding prescription medication use.	High agreement was observed for prescription drugs used for longer duration (chronic), like insulin and its analogues (κ=0.78) and cardiovascular drugs (κ=0.84).	Nice study design. The findings of the study indicated higher recall rates for prescription drugs used for long duration or for chronic conditions, and lower recall rates for drugs used occasionally.
Sjahid et al. (1998) [66]	Cohort of patients from the Rotterdam elderly Study (n=3,365)	Pharmacy data	Self-reported information regarding cardiovascular drugs	High overall agreement was observed (80.6%). Highest agreement was noted for β-adrenoceptor class (κ=0.97).	The findings indicated that agreement was good for drugs that are used repetitively /regularly, than those that are used intermittently. Also in this dataset, the authors noted that agreement was good for drugs that are 'prescription only' over those that are also available over-the-

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
					counter.
Curtis et al. (2006) [73]	Cohort of patients using medication for chronic glucocorticoid (n=2,363)	Pharmacy data from a national MCO	Self-reported information regarding use of four osteoporosis medications (alendronate, risedronate, calcitonin, raloxifene)	Substantial to moderate agreement was found for the four medications, highest being for alendronate (κ=0.80).	Good study design. The authors observed high agreement for medications with characteristic dosing instructions than for those that have less exclusive dosing instructions.
Løkkegaard et al. (2004) [72]	Cohort of Danish nurses (n=2,666)	National level prescription reimbursement database	Self-reported information about hormone replacement therapy (HRT)	Overall high accuracy rate was found for self-reported HRT use. For current HRT use sensitivity was 78.4% and specificity was 98.4%.	Good study design. Study sample used for the study (nurses) might have contributed towards higher degree of agreement.
Brown et al. (2007) [67]	Cohort of patients from the Adverse Childhood Experiences (ACE) study (n=4,308)	Pharmacy claims data	Self-reported information about exposure to lipid-lowering drugs	Good agreement was observed for self-reported lipids lowering medication use (κ=0.67); 96% cases were concordant.	--
STUDIES CONDUCTED AMOMG PREGNANT WOMEN					
Bryant et al. (1989) [30]	Cohort of pregnant women (n= 202)	Medical records	Short term illnesses and medication use during pregnancy	In general lower level of agreement was observed for most conditions and medication use. Highest agreement was noted for prescription medication use (κ=0.48).	Used two groups of women: still pregnant and postpartum. Comparison of results among these two groups provided nice insight into issues related to recall of events and minimized

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
					biases related to it.
Webb et al. (2003) [32]	Cohort of pregnant women (n= 74)	Laboratory urine cotinine level results	Self-reported smoking behavior	The results of the study showed high level of disagreement. While laboratory results showed 63% of the women smoked during pregnancy, only 25% of the women gave correct self-reported information about smoking.	Use of urine cotinine levels provided a good measure to validate self-reported information. Also the population sample used (low-income) provided good insight on the characteristics of pregnant women in this population, thereby providing knowledge about how to effectively plan smoking cessation programs in such a population.
Ford et al. (1997) [33]	Cohort of pregnant women (n= 4, 857)	Laboratory serum cotinine level results, obstetrics records	Self-reported smoking behavior	There was substantial level of discrepancy between self-reported information about smoking behavior and laboratory serum cotinine level results. Only 19.2% of the women gave accurate information about smoking in the first trimester, whereas form laboratory results it was found 31.3% of the women smoked during first trimester.	Identical results for level of agreement in self-reported information and obstetric records (which were inconsistent with lab results) to compare the revealed that self-reported information about smoking status provided by pregnant women to their physicians may be inaccurate, and should be validated by a biochemical test.
Ernhart et al. (1988) [34]	Cohort of pregnant	Retrospectively collected	Self-reported information regarding	There was significant discrepancy between the	Long recall interval (5 years) might have

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	women (n= 238)	information regarding alcohol consumption during pregnancy	alcohol consumption during pregnancy	two data sources. More women reported positive smoking during pregnancy when asked retrospectively than when they were asked during pregnancy.	contributed towers lowering the level of agreement between the two data sources.
Klebanoff et al. (2001) [35]	Cohort of pregnant women (n= 105)	Laboratory serum/urine cotinine level results	Self-reported smoking behavior	Good level of agreement was observed. Concordance was found for self-reported information from 84.6% women when compared with their laboratory serum/urine cotinine levels.	The study sample wad derived from another study for calcium intake and pre-eclampsia and was not focused on smoking behavior, which might have contributed to moderate agreement between the two sources of data.
Klebanoff et al. (1998) [42]	Cohort of pregnant women (n=452)	Laboratory serum cotinine level results	Self-reported smoking behavior	High level of agreement was found for data sources (κ=0.83).	--
Lester et al. (2001) [107]	Cohort of women who had just given birth (n=8,527)	Newborn baby's meconium analysis results	Self-reported drug use	Concordant results were obtained in 66 % cases.	Use of meconium analysis results is a novel method to validate self-reported information about drug use.
Jacobson et al. (1991) [81]	Cohort of pregnant women (n=361)	Patients' Michigan Alcoholism Screening Test (MAST)	Self-reported information about alcohol and drug use during pregnancy (collected twice prenatally and 13	Women reported higher alcohol and drug use in their retrospective self-reports, than when they were asked during their pregnancy. Also, these	Use of MAST scores for validating self-reported alcohol and drug use is unique. These measures also revealed maternal characteristics for high

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
		scores	months after birth of their babies)	women had higher MAST scores.	alcohol consumption during pregnancy like, depression and a history of alcohol abuse.
Olesen et al. (2001) [40]	Cohort of pregnant women enrolled in the DNBC (Danish National Birth Cohort) survey (n=2,041)	North Jutland Prescription database	Self-reported use of prescription medication use	In general, higher recall rates were observed for medications taken recently and those taken for longer duration. For example, recall rate was 43% for medications taken 120 days before the interview and 49% for medication taken 60 days before the interview.	The results of the study suggested higher recall rates for chronically used drugs. The authors were able to describe potential biases that could have occurred due to the data sources used for comparison.
Fox et al. (1989) [80]	Cohort of pregnant women enrolled in a randomized clinical trial of a smoking cessation program (n=700)	Saliva thiocyanate levels	Self-reported smoking status and alcohol consumption during pregnancy.	Identical rates of recall were observed for both the treatment group and the control group. The level of agreement of smoking status for the treatment group was κ=0.61 and κ=0.57 for the control group. For alcohol use it was κ=0.52 for the treatment group and κ=0.55 for the control group.	Self-reported information was collected twice. The results reveal that pregnant women provide similar information about smoking status at any time interval during their pregnancy.
Hessol et al. (2004) [79]	Cohort of pregnant	Medical records	Behavioral factors (alcohol use, tobacco	Lower agreement for alcohol use during	Good study design. The study illustrates the

Author (year)	Population and Sample size	Gold-standard	Conditions/Features Evaluated	Results/ Key Findings	Comments
	women in the Latina Health Project (n= 321)		use, use of prenatal vitamins) and medical factors (anemia, gestational diabetes, hypertension).	pregnancy (κ=0.37) and prenatal vitamin use (κ=0.09) was observed. Agreement for gestational diabetes (κ=0.83) and hypertension (κ=0.68) was the highest.	influence of language on recall accuracy.
Britton et al. (2004) [77]	Cohort of pregnant women enrolled in a smoking cessation program (n= 94)	Laboratory urine cotinine level results	Self-reported smoking status	Disconcordance was found in 34.7% of cases where active (current) smokers reported no smoking, and for 10.4% of women who reported active smoking but their urine cotinine levels.	Awareness to quit smoking among pregnant women enrolled in the smoking cessation program might have added to increase in agreement.

APPENDIX 2: ‘SMART’ STUDY QUESTIONNAIRE

GENERAL INFORMATION

1. Date of interview: ___ /___ /___ (month/day/year)
2. Location of interview: _____
3. Prenatal care provider's last name: _____
4. Examiner's last name: _____
5. Patient's phone number: _____

DEMOGRAPHIC / LIFESTYLE INFORMATION

6. How old are you? _____ (years)
7. What is your marital status now?

   [ ] Single, never married
   [ ] Married, living with spouse
   [ ] Not married, but living with partner
   [ ] Separated from spouse
   [ ] Divorced
   [ ] Widowed

8. Are you Hispanic, Latino or of Spanish descent? [ ] Yes [ ] No
9. How do you describe yourself: (check all that apply)

   [ ] White, non-Hispanic or [ ] White, Hispanic
   [ ] Black or African American
   [ ] American Indian or Alaskan Native - Please specify [ ] tribe or [ ] pueblo
   [ ] Asian or Asian American or Pacific Islander
   [ ] Some other group(s) – please specify: _____
   [ ] Prefer not to report

10. What is the highest level in school you have completed?

    [ ] Less than high school graduate
    [ ] High school graduate or GED
    [ ] Some college or vocational school
    [ ] College degree
    [ ] Masters, doctorate or professional degree

11. What is your health insurance status?

[ ] No insurance

[ ] Employer-based insurance

[ ] Self-purchased insurance

[ ] Medicaid

[ ] Other public insurances ([ ] Indian Health Service, [ ] VA, [ ] First Choice, UNM/UNMCARE)

11a. Does your insurance cover prescription drugs? [ ] Yes [ ] No

12. Were you born in the Unites States? [ ] Yes [ ] No

If 'Yes', go to question 13. If 'No', please answer questions 12a and 12b.

12a. Did you move to the United States: [ ] With your parents when you were a child

[ ] When you were an adult (≥18 years old)

12b. How long have you lived in the United States: _____ years

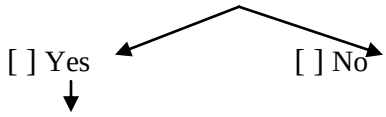
13. What language do you mostly use at home?

[ ] English

[ ] Spanish

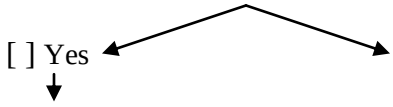
[ ] Some other language – specify:_____

14. Do you currently smoke cigarettes or use tobacco?



14a. How many cigarettes do you usually smoke in one day? _____

14b. Have you smoked >100 cigarettes in your life?



14c. When did you stop smoking? * *Go to the next question*

[ ] Before I became pregnant

[ ] After I realized that I was pregnant

15. Have you ever drank alcohol in your life (e.g., beer, wine, hard liquor, mixed drinks)?
[ ] Yes [ ] No

If 'yes,' continue to questions 15a and 15b. If 'no,' continue to 21a.

15a. How many drinks does it take before you begin to feel the first effect of alcohol? _____

15b. How many drinks typically can you hold before passing out or falling asleep? _____

a. What was the first day of your last menstrual period ___/___/___ (mm/dd/yy)?

I would like you to think back to that period and tell me about your drinking at that time.

16. During a month or so around your last menstrual period before you got pregnant, how many times did you drink **4 or more drinks** on one occasion? _____

Now I want you to think of <u>12 months before you got pregnant</u> (a year prior to your LMP)

17. During the year before you got pregnant, did close friends or relatives worry or complain about your drinking habits?

[ ] Yes [ ] No

18. During the year before you got pregnant, did you ever take a drink first thing in the morning to get yourself going?

[ ] Yes [ ] No

19. During the year before you got pregnant, did a friend or family member tell you about things you said or did while you were drinking that you could not remember?

[ ] Yes [ ] No

20. During the year before you got pregnant, did you feel you need to cut down on your drinking?

[ ] Yes [ ] No

[TWEAK High: _____; TWEAK Hold: _____]

<u>MEDICAL AND REPRODUCTIVE HEALTH</u>

21a. What was your pre-pregnancy weight? _____ pounds

21b. What was your pre-pregnancy height? _____ feet/inches

[Researcher Calculated BMI: _____]

22. Do you have a medical condition or problem that requires ongoing, periodic, or occasional treatment?
[ ] Yes [ ] No

22a. If yes, check all that apply:

[ ] Hypertension (high blood pressure)
[ ] Diabetes: [ ] Gestational [ ] Type I [ ] Type II
[ ] Seizure disorder (i.e., epilepsy)
[ ] Thyroid disorder
[ ] Asthma or allergies
[ ] Cancer
[ ] Depression
[ ] Anxiety
[ ] Migraine headaches
[ ] Rheumatoid arthritis
[ ] Heart disease
[ ] Hepatitis
[ ] Other(s) problem - specify: _____

If 'Yes' to diabetes, please answer questions 23 and 24. If 'No', skip to question 26.
If 'Yes' to asthma, please answer question 25. If 'No,' skip to question 26.

23. Have you ever had gestational diabetes?

[ ] Yes, in a previous pregnancy only
[ ] Yes, in the current pregnancy only
[ ] Yes, in a previous pregnancy and in the current pregnancy
[ ] No, never had gestational diabetes
[ ] No, never been pregnant before

24. How likely do you think uncontrolled high blood sugar could harm your developing baby by causing birth defects or other serious health problems? (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

25. How likely do you think asthma exacerbations requiring hospitalization or unscheduled clinic visits could harm your developing baby? (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

26. Did you plan to get pregnant with this child?

[ ] Yes [ ] No, not now [ ] No, not at any time

27. Were you or your partner doing anything to try to prevent becoming pregnant with this child?

[ ] Yes [ ] No

27a. If *Yes*, which method were you using?

[ ] Condoms [ ] Diaphragm [ ] Birth control pills

[ ] Withdrawal [ ] IUD [ ] Rhythm

[ ] Depo Provera, Implanon or Norplant [ ] Other: _____

28. Did you take any fertility drugs to help you get pregnant with this child, like Clomid, Metrodin, Fertinex, or Pergonal?

[ ] Yes [ ] No

28a. If *Yes*, which drugs did you use? _____

29. Have you or members of your immediate family (mother or sisters) or the immediate family of your baby's father had any babies with birth defects (including babies that might not have survived)?

[ ] Yes [ ] No

** If 'No', go to question 30. If 'Yes', please specify:*

Down syndrome	[ ] Yes	[ ] No
Cleft lip or palate	[ ] Yes	[ ] No
Neural tube defect	[ ] Yes	[ ] No
Cystic fibrosis	[ ] Yes	[ ] No
Heart defect	[ ] Yes	[ ] No
Other	[ ] Yes	[ ] No

If "Yes," please specify:_____

30. What was the **first day** of your last menstrual period?___ ___ / ___ ___ /___ ___ *(mm / dd / yy)*

31. What is the date your baby is due to be born: ___ ___ / ___ ___ / ___ ___ *(mm / dd / yy)*

31a. What is the gestational age of your baby? _____ weeks

31b. How was your due date estimated? By:

[ ] Last menstrual period
[ ] Ultrasound
[ ] Physical exam

32. How many times (including this pregnancy) have you been pregnant?
_____

If this is the 1st pregnancy put "1"for q. 32 and "0" for questions 33-37 & skip to question 37.

33. How many live-born children have you had? _____
If no live-born children or this is the first pregnancy, then put "0"

34. Have you ever had a miscarriage (<20 wk of gestation). If yes, how many?_____
If never had a miscarriage, put "0"

35. Have you ever had a stillborn child (≥ 20 wk of gestation). If yes, how many?_____
If never had a stillborn child, put "0"

36. Have you ever had a pregnancy terminated? If yes, how many?_____
If never had a termination, put "0"

37. Have you ever had an ectopic pregnancy. If yes, how many?_____
If never had an ectopic pregnancy, put "0"

38. For this pregnancy, how many weeks after your last menstrual period did you first think you were pregnant? _____

39. For this pregnancy, how many weeks after your last menstrual period did you first go to see a doctor or other health care provider or go to the clinic for prenatal care?
_____

40a. Have you had any complications in this pregnancy so far?
(* *Please check yes or no for each complication)*

- Bleeding [ ] Yes [ ] No
- High blood pressure [ ] Yes [ ] No
- Diabetes [ ] Yes [ ] No
- Other [ ] Yes [ ] No

40b. If "other", please specify: _____

41. Have you experienced morning sickness during this pregnancy? [ ] Yes [ ] No

USE OF MEDICATIONS AND SUPPLEMENTS DURING PREGNANCY

42. Did you take a multivitamin regularly (4 times a week or more) during the month before your last menstrual period?

[ ] Yes [ ] No

43. Have you taken any **VITAMINS** regularly (4 times/week or more) since you became pregnant?

43a. [ ] Yes, multivitamins [ ] Yes, a single vitamin [ ] No

If 'Yes,' answer questions 43b-43e.

43b. [ ] Prescription [ ] OTC

43c. [ ] Brand name: _____

43d. When did you start taking vitamins?

_____ (mm/dd/yy) _____ (gestational weeks)

43e. How many days during the last week did you take vitamins? _____ (days/week)

44a. Have you taken any **DIETARY SUPPLEMENTS** (including iron supplements) or **HERBAL PRODUCTS** on a regular basis since your last menstrual period?

[ ] Yes [ ] No

44b. If 'Yes' to herbal products, please specify: [ ] Herbs

[ ] Tablets or capsules

[ ] Teas

[ ] Other: _____

44c. How often do you take them? [ ] Regularly: _____ times per _____ or [ ] When I feel sick

Please specify any other dietary supplements or products and reason for taking it:

Product 1:_____ Reason/Condition: _____

Product 2: _____Reason/Condition: _____

Product 3: _____Reason/Condition: _____

44d. Have you had any cravings for non-food items or really "strange" foods?

[ ] Yes [ ] No

If 'yes' what did you crave, do you eat it, and how often do you eat it?

Item 1: _____ Eat it? [ ] Yes [ ] No; How often?

Item 2: _____ Eat it? [ ] Yes [ ] No; How often?

Item 3: _____ Eat it? [ ] Yes [ ] No; How often?

45. Have you ever taken any recreational drugs?
[ ] Yes [ ] No

If 'Yes' please specify the recreational drug name(s) and when it was used:

Check if taken:

[ ] Marijuana/Hashish: [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

[ ] Heroin: [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

Have you gone through methadone treatment?
[ ] Never
[ ] Completed treatment before pregnancy
[ ] Undergoing treatment during current pregnancy

[ ] Cocaine/Crack: [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

[ ] Inhalants (glue, solvent): [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

[ ] Methamphetamines: [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

[ ] Other: _____ [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

[ ] Other: _____ [ ] Before pregnancy
[ ] 1 month prior to LMP or during this pregnancy

46. Did you discuss the safety of medications in pregnancy with any health care provider (physician, nurse-midwife, physician assistant, or pharmacist)?

[ ] Yes [ ] No

47a. Have you had any vaccinations since your last menstrual period?

[ ] Yes [ ] No

47b. If *Yes* to vaccinations, please specify: [ ] Flu

[ ] Other: _____

48. Have you taken any **medications PRESCRIBED by your doctor** or any other health care provider since your last menstrual period, even if you stopped taking them once you knew you were pregnant?

[ ] Yes [ ] No

If 'Yes' please specify the medication name, reason for taking it, and <u>your perception</u> of how likely it is that this medication might be harmful for your baby if taken during pregnancy:

a. **Medication1**:_____ Indication: _____

How likely do you think it is that this medication could harm your developing baby by causing birth defects or other serious health problems: (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

b. **Medication2**:_____ Indication: _____

How likely is it that this medication could harm your developing baby by causing birth defects or other serious health problems: (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

c. **Medication 3**:_____ Indication: _____

How likely is it that this medication could harm your developing baby by causing birth defects or other serious health problems: (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

d. **Medication 4**:_____ Indication: _____

How likely is it that this medication could harm your developing baby by causing birth defects or other serious health problems: (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

e. **Medication 5**:_____ Indication: _____

How likely is it that this medication could harm your developing baby by causing birth defects or other serious health problems: (circle one number)

1	2	3	4	5
Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to cause harm

49. During this pregnancy, did you take any **OVER-THE-COUNTER MEDICATIONS** (sold without prescription)?

[ ] Yes [ ] No

Check all medications that you have <u>actually taken</u> since your last menstrual period, even if you stopped taking them once you knew you were pregnant. Then for medications you took since pregnancy, please specify <u>your perception</u> of how likely each medication is to cause birth defects or other problems for your baby.

Pain/Fever Medications:

<u>Rate all medications</u>**:** <u>Check if taken:</u>	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Acetaminophen (Tylenol)	1	2	3	4	5
[ ] Aspirin	1	2	3	4	5
[ ] Ibuprofen (Advil, Motrin)	1	2	3	4	5
[ ] Ketoprofen (Orudis)	1	2	3	4	5
[ ] Naproxen (Aleve)	1	2	3	4	5
[ ] Other medication – specify:	1	2	3	4	5

_____

Nasal Decongestants, Allergy, Cough Medications:

<u>Rate all medications</u>**:** <u>Check if taken:</u>	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Chlorpheniramine (Chlor-Trimeton)	1	2	3	4	5
[ ] Benadryl	1	2	3	4	5
[ ] Pseudoephedrine (Sudafed)	1	2	3	4	5

[ ] Claritin, Zyrtec	1	2	3	4	5
[ ] Other medication –	1	2	3	4	5

specify: _____

Antidiarrheal Medications:

Rate all medications: Check if taken:	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Kaopectate, Pepto Bismol	1	2	3	4	5
[ ] Loperamide (Imodium)	1	2	3	4	5
[ ] Other medication –	1	2	3	4	5

specify: _____

Heartburn, Dyspepsia, Antiemetic, Laxative Medications:

Rate all medications: Check if taken:	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Maalox, Mylanta Gas	1	2	3	4	5
[ ] Tums	1	2	3	4	5
[ ] Tagamet, Zantac, Axid, Pepcid	1	2	3	4	5
[ ] Colace	1	2	3	4	5
[ ] Correctol, Dulcolax, Ex-Lax	1	2	3	4	5
[ ] Senna, fiber products	1	2	3	4	5
[ ] Unisom	1	2	3	4	5
[ ] Other medication –	1	2	3	4	5

specify: _____

Antifungal Medications (taken for vaginal yeast infection or thrash):

Rate all medications: Check if taken:	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Vaginal cream or suppositories (Monistat, Vagistat, Femstat, Lotrim)	1	2	3	4	5

[ ] Other medication –	1	2	3	4	5

specify: _____

Nicotine Replacement Therapy (for smoking cessation):

Rate all medications: Check if taken:	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Nicotine gum, spray or inhaler	1	2	3	4	5
[ ] Nicotine patch	1	2	3	4	5
[ ] Other medication –	1	2	3	4	5

specify: _____

Other over-the-counter medications you have taken while pregnant:

Rate all medications: Check if taken:	Not at all likely to harm	Unlikely to cause harm	Somewhat likely to harm	Likely to harm	Very likely to harm
[ ] Other medication –	1	2	3	4	5

specify: _____

[ ] Other medication –	1	2	3	4	5

specify: _____

[ ] Other medication –	1	2	3	4	5

specify: _____

50. If you took prescription medications regularly before you got pregnant, did you change the use of these medications when you realized you are pregnant?

[ ] Did not take prescription medications regularly before pregnancy

[ ] Discontinued the use upon recognition of pregnancy.

Medication_____

[ ] Decreased the use (dose or frequency).

Medication: _____

[ ] Increased the use.

Medication: _____

[ ] Stayed the same, continued without any change.

Medication: _____

50a. If you changed the use of a medication upon recognition of pregnancy, why?

[ ] Provider recommendation

[ ] Family or friend suggestion

[ ] Self-initiated

[ ] Financial constraints

[ ] Other: _____

Now I'm going to ask you about your thoughts about medication use during pregnancy in general. Please check the answer you think is the most appropriate for each question.

51. If a woman plans a pregnancy or finds out that she is currently pregnant, she should:

[ ] Stop taking all medications immediately to protect the baby

[ ] Continue taking only those medications that are absolutely necessary and check with her doctor to see if the medications are safe for the baby

[ ] Continue taking necessary medications but reduce the dose or the number of days you take them to limit the amount that gets to the baby

[ ] Continue with all medications as needed since medications are safe for the baby

52. When a woman uses medications regularly during pregnancy, how often can medications
cause birth defects?

[ ]	[ ]	[ ]	[ ]	[ ]
Never	Sometimes	Often	Very Often	Always

53. Which statement best describes your view about women drinking alcohol during pregnancy?

[ ] Pregnant women should abstain from drinking any alcohol (even small amounts) during pregnancy.

[ ] It is OK for a woman to have an occasional drink during pregnancy as long as it is not more often than once a week.

[ ] It is OK for a woman to have an occasional drink during pregnancy as long as it is not more often than one drink per day.

[ ] It is OK for a woman to have an occasional drink during pregnancy as long as it is not more often than two drinks per day.

[ ] It is OK for a woman to drink during pregnancy as long as she does not drink hard liquor (i.e., vodka, whiskey, brandy) but only drinks wine or beer.

54. During your current pregnancy, have you ever asked anyone about the safety of medications you are taking for your baby?
[ ] Yes [ ] No

54a. If yes, check any individuals who you have asked a question about the safety of any medications for your baby: (*Check all that apply to you*)

[ ] Your primary care doctor or provider

[ ] Your OB/GYN doctor or midwife

[ ] A pharmacist

[ ] A member of your family, spouse

[ ] A friend, partner

[ ] Other – specify: _____

[ ] Any other heath care provider

55. Please check any sources below in which you have looked for information about the safety of medications for your baby? (*Check all that apply to you*)

[ ] I have never looked at any of these sources about the safety of medications for my baby

[ ] An internet web site(s). Specify: _____

[ ] A book.

[ ] A magazine

[ ] Pregnancy information telephone service/hotline (i.e., OTIS, Nurse Advisory Line)

[ ] Other – Specify:_____

[ ] I have not had any questions about the safety of medications for my baby and have not looked at any of these sources.

[ ] Clinic pamphlet or brochure

56. NOTES/COMMENTS:

_____

_____

_____

_____

_____

APPENDIX 3: SAS-MACRO USED FOR ESTIMATING PREVALENCE-ADJUSTED AND BIAS-ADJUSTED KAPPA (PABAK)

```
/*GENEARTE A WINDOW TO INPUT RATING RESPONSES*/

  %window kap color=white
#2 @33 'Rater Agreement' attr=(highlight,underline) color=blue
#4 @34 '2 Raters Only' attr=highlight color=blue
#7 @19 'Enter the counts a, b, c, and d in the table below:' attr=highlight
color=blue #8 @25 '(Use the TAB key to jump to next cell)' color=blue
#11 @16 ' RATER B '
#12 @38 'YES' attr = highlight @68 'NO' attr = highlight
#13 @18 '_____________________________________________________________'
#15 @15 'YES' attr = highlight @22 'Cell A' @34 a 10 attr = underline
required=yes
#15 @52 'Cell B' @64 b 10 attr = underline required=yes
#16              @3              'RATER              A'              @18
'_____________________________________________'
#18 @16 'NO' attr = highlight @22 'Cell C' @34 c 10 attr = underline
required=yes
#18 @52 'Cell D' @64 d 10 attr = underline required=yes
#19 @18 '_____________________________________________________________'
#23 @22 'YES or NO indicate the dichotomous responses for each rater'
#30 @33 'Press ENTER to continue' attr=highlight ;
%macro def ;
%let a = ;
%let b = ;
%let c = ;
%let d = ;
%display kap ;
%mend def ;
%def ;data one;
set two;
if mix(var1, var2) > 0 then do;

/*DEFINE FORMATS*/

proc format ;
value rating
0 = "poor"
1 = "slight"
2 = "fair"
3 = "moderate"
4 = "substantial"
5 = "almost perfect"
6 = "cannot calculate kappa"
;
value rs
1 = "yes"
2 = "no"
;
run ;
```

```
/*CALCULATE THE MEASURES OF AGREEMENT*/


data calcs ;
a = &a ;
b = &b ;
c = &c ;
d = &d ;
N = a+b+c+d ;

po = (a+d)/N ;
pe = ((a+c)*(a+b)+(b+d)*(c+d))/N**2 ;
ppos = (2*a)/(N+a-d) ;
pneg = (2*d)/(N-a+d) ;


pi = (a-d)/N ;
bi = (b-c)/N ;
pabak = 2*po-1 ;

kappa = (po-pe)/(1-pe) ;
q = ((a/N)*(1-(((a+b)/N)+((a+c)/N))*(1-kappa))**2)+((d/N)*
       (1-(((c+d)/N)+((b+d)/N))*(1-kappa))**2);
r = ((1-kappa)**2)*((b/N)*(((a+c)/N)+((c+d)/N))**2+(c/N)*
(((b+d)/N)+((a+b)/N))**2) ;
s = (kappa - pe*(1-kappa))**2 ;
*Asymptotic standard error ;
se_kappa = sqrt((q+r-s)/(N*(1-pe)**2)) ;
LL_95_CI = kappa-1.96*se_kappa ;
if LL_95_CI < -1.00 then LL_95_CI = -1.00 ;
UL_95_CI = kappa+1.96*se_kappa;
if UL_95_CI > 1  then UL_95_CI = 1.00 ;
se_kappa_null =sqrt((1/(N*(1-pe)**2))*(pe+(pe**2)-
((((a+b)/N)*((a+c)/N)*(((a+b)/N)+((a+c)/N))+((c+d)/N)*((b+d)/N)*(((c+d)/
N)+((b+d)/N))))) ;
z = kappa/se_kappa_null ;
p = 1 - cdf('Normal',z,0,1) ;
label se_kappa = "Kappa Std. Error" ;
label LL_95_CI = "95% CI Lower Limit" ;
label UL_95_CI = "95% CI Upper Limit" ;
label se_kappa_null = "Kappa Std. Error (Under Ho)" ;
label z = "Z (Under Ho:Kap=0)" ;
label p = "One sided p-value (Under Ho:Kap=0)" ;
label po = "Observed Agreement (Po)" ;
label pe = "Expected Agreement (Pe)" ;
label ppos = "Positive Agreement (Ppos)" ;
label pneg = "Negative Agreement (Pneg)" ;
label pi = "Prevalence Index" ;
label bi = "Bias Index" ;
label kappa = "Kappa" ;
label pabak = "PABAK" ;
```

```
strength = 0 ;
if kappa gt 0.00 and kappa le 0.20 then strength = 1 ;
if kappa gt 0.20 and kappa le 0.40 then strength = 2 ;
if kappa gt 0.40 and kappa le 0.60 then strength = 3 ;
if kappa gt 0.60 and kappa le 0.80 then strength = 4 ;
if kappa gt 0.80 and kappa le 1.00 then strength = 5 ;
if kappa = . then strength = 6 ;
format strength rating. ;
label strength = "Strength of Agreement" ;
run ;

proc print data = calcs label noobs ;
var kappa strength se_kappa LL_95_CI UL_95_CI se_kappa_null z p po pe ppos pneg
            pi bi pabak ;
format po pe ppos pneg pi bi kappa se_kappa LL_95_CI UL_95_CI se_kappa_null p
pabak 6.4 z 4.2 ;
run ;
```

REFERENCES

1. Malm, H., et al., *Prescription drugs during pregnancy and lactation--a Finnish register-based study.* Drug Safety, 2003. **59**(2): p. 127-33.
2. Andrade, S.E., et al., *Prescription drug use in pregnancy.* Am J Obstet Gynecol, 2004. **191**(2): p. 398-407.
3. Riley, E.H., et al., *Correlates of prescription drug use during pregnancy.* Journal of Women's Health, 2005. **14**(5): p. 401-09.
4. Piper, J.M., C. Baum, and D.L. Kennedy, *Prescription drug use before and during pregnancy in a Medicaid population.* Am J Obstet Gynecol, 1987. **157**(1): p. 148-56.
5. Glover, D.D., et al., *Prescription, over-the-counter, and herbal medicine use in a rural, obstetric population.* Am J Obstet Gynecol, 2003. **188**(4): p. 1039-45.
6. Lacroix, I., et al., *Prescription of drugs during pregnancy in France.* The Lancet, 2000. **356**(9243): p. 1735-36.
7. Henshaw, S.K., *Unintended pregnancy in the United States.* Fam Plann Perspect, 1998. **30**(24): p. 24-46.
8. Briggs GG, Freeman RK, and Yaffe SJ, *Drugs in pregnancy and lactation*. Vol. 12. 2008. 440.
9. Koren, G., A. Pastuszak, and S. Ito, *Drugs in pregnancy.* N Engl J Med, 1998. **338**(16): p. 1128-1137.
10. Howland, R.H., *Categorizing the safety of medications during pregnancy and lactation.* Journal of Psychosocial Nursing, 2009. **47**(4): p. 17-20.

11. *United States Food and Drug Administration (FDA) Proposed Rules.* Federal Register, 2008. **73**(104): p. 30331-30368.

12. Kriegsman, D.M.W., et al., *Self-reports and general practitioner information on the presence of chronic diseases in community dwelling elderly A study on the accuracy of patients' self-reports and on determinants of inaccuracy.* J Clin Epidemiol, 1996. **49**(12): p. 1407-17.

13. Haapanen, N., et al., *Agreement between questionnaire data and medical records of chronic diseases in middle-aged and elderly Finnish men and women.* Am J Epidemiol, 1997. **145**(8): p. 762-69.

14. Bush, T.L., et al., *Self-report and medical record report agreement of selected medical conditions in the elderly.* Am J Public Health, 1989. **79**(11): p. 1554-56.

15. Okura, Y., et al., *Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure.* J Clin Epidemiol, 2004. **57**(10): p. 1096-1103.

16. Paganini-Hill, A. and R.K. Ross, *Reliability of recall of drug usage and other health-related information.* Am J Epidemiol, 1982. **116**(1): p. 114-22.

17. Smith, B., et al., *Challenges of self-reported medical conditions and electronic medical records among members of a large military cohort.* BMC Medical Research Methodology, 2008. **8**(1): p. 37.

18. Beckles, G.L.A., et al., *Agreement Between Self-Reports and Medical Records Was Only Fair in a Cross-Sectional Study of Performance of Annual Eye Examinations Among Adults With Diabetes in Managed Care.* Med Care, 2007. **45**(9): p. 876-83.

19. Centers for Disease Control and Prevention (CDC). Behavioral Risk Factor Surveillance System Survey Data: U.S. Department of Health and Human Services. 2009 [cited 2009 September 9].

20. Simpson, C.F., et al., *Agreement Between Self-Report of Disease Diagnoses and Medical Record Validation in Disabled Older Women: Factors That Modify Agreement.* J Am Geriatr Soc, 2004. **52**(1): p. 123-27.

21. *National Patient Safety Goals*. The Joint Commission, 2009 [cited 2009 September 9].

22. Cornish, P.L., et al., *Unintended medication discrepancies at the time of hospital admission.* Arch Intern Med, 2005. **165**(4): p. 424-429.

23. Schnipper, J., C. Hamann, and C. Ndumele, *Effect of an electronic medication reconciliation application and process redesign on potential adverse drug events: a clsuter randomized trial.* Arch Intern Med, 2009. **169**: p. 771-780.

24. Lo, W.Y. and J.M.M.D. Friedman, *Teratogenicity of Recently Introduced Medications in Human Pregnancy.* Obstet Gynecol, 2002. **100**(3): p. 465-73.

25. Lagoy, C.T., et al., *Report from the CDC: Medication Use during Pregnancy and Lactation: An Urgent Call for Public Health Action.* Journal of Women's Health, 2005. **14**(2): p. 104-09.

26. Tisnado, D.M., et al., *What is the concordance between the medical record and patient self-report as data sources for ambulatory care?* Med Care, 2006. **44**(2): p. 132-40.

27. Harlow, S.D. and M.S. Linet, *Agreement between questionnaire data and medical records: the evidence for accuracy of recall.* Am J Epidemiol, 1989. **129**(2): p. 233-48.

28. West, S.L., et al., *Recall accuracy for prescription medications: self-report compared with database information.* Am J Epidemiol, 1995. **142**(10): p. 1103-12.

29. Van den Brandt, P.A., et al., *Comparison of questionnaire information and pharmacy data on drug use.* Pharm Weekbl [Sci], 1991. **13**(2): p. 91-96.

30. Bryant, H.E., N. Visser, and E.J. Love, *Records, recall loss, and recall bias in pregnancy: a comparison of interview and medical records data of pregnant and postnatal women.* Am J Public Health, 1989. **79**(1): p. 78-80.

31. Yawn, B.P., V.J. Suman, and S.J. Jacobsen, *Maternal recall of distant pregnancy events.* J Clin Epidemiol, 1998. **51**(5): p. 399-405.

32. Webb, D.A., et al., *The Discrepancy Between Self-Reported Smoking Status and Urine Cotinine Levels Among Women Enrolled in Prenatal Care at Four Publicly Funded Clinical Sites.* J Public Health Management Practice, 2003. **9**(4): p. 322-25.

33. Ford, R.P., et al., *Smoking during pregnancy: how reliable are maternal self reports in New Zealand?* J Epidemiol Community Health, 1997. **51**(3): p. 246-51.

34. Ernhart, C.B., et al., *Underreporting of alcohol use in pregnancy.* Alcohol Clin Exp Res, 1988. **12**(4): p. 506-11.

35. Klebanoff, M.A., et al., *Accuracy of self-reported cigarette smoking among pregnant women in the 1990s.* Paediatric and Perinatal Epidemiology, 2001. **15**(2): p. 140-43.

36. Westbrook, J.I., et al., *Agreement between medical record data and patients' accounts of their medical history and treatment for dyspepsia.* J Clin Epidemiol, 1998. **51**(3): p. 237-44.

37. Caplan, L.S., et al., *Validity of women's self-reports of cancer screening test utilization in a managed care population.* Cancer Epidemiol Biomarkers Prev, 2003. **12**(11): p. 1182-87.

38. Sauver, J.L.S., et al., *Agreement between patient reports of cardiovascular disease and patient medical records.* Mayo Clin Proc, 2005. **80**(2): p. 203-10.

39. Glintborg, B., et al., *Are patients reliable when self-reporting medication use? Validation of structured drug interviews and home visits by drug analysis and prescription data in acutely hospitalized patients.* J Clin Pharmacol, 2007. **47**(11): p. 1440-49.

40. Olesen, C., et al., *Do pregnant women report use of dispensed medications?* Epidemiology, 2001. **12**(5): p. 497-501.

41. Camilli, A.E., L.F. McElroy, and K.L. Reed, *Smoking and pregnancy: a comparison of Mexican-American and non-Hispanic white women.* Obstet Gynecol, 1994. **84**(6): p. 1033-37.

42. Klebanoff, M.A., et al., *Serum cotinine concentration and self-reported smoking during pregnancy.* Am J Epidemiol, 1998. **148**(3): p. 259-62.

43. Headley, J., et al., *Medication use during pregnancy: data from the Avon Longitudinal Study of Parents and Children.* Euro J Clin Pharmacol, 2004. **60**(5): p. 355-61.

44. Ferrante, J.M., et al., *Self-report versus Medical Records for Assessing Cancer-Preventive Services Delivery.* Cancer Epidemiol Biomarkers Prev, 2008. **17**(11): p. 2987-94.

45. Bergmann, M.M., et al., *Validity of self-reported Cancers in a Propsective Cohort Study in Comparison with Data from State Cancer Registries.* Am J Epidemiol, 1998. **147**(6): p. 556-62.

46. Hall, H.I., et al., *Testing for prostate and colorectal cancer: comparison of self-report and medical record audit.* Prev Med, 2004. **39**(1): p. 27-35.

47. Mukerji, S.S., et al., *Comorbidities in head and neck cancer: Agreement between self-report and chart review.* Otolaryngol Head Neck Surg, 2007. **136**(4): p. 536-42.

48. Skinner, K.M., et al., *Concordance Between Respondent Self-reports and Medical Records for Chronic Conditions: Experience From the Veterans Health Study.* J Ambul Care Manage, 2005. **28**(2): p. 102-10.

49. Goebeler, S., M. Jylhb, and A. Hervonen, *Self-reported medical history and self-rated health at age 90. Agreement with medical records.* Aging Clin Exp Res, 2007. **19**(3): p. 213-19.

50. Solomon, D.H., et al., *Agreement between patient report and medical record review for medications used for rheumatoid arthritis: The accuracy of self-*

reported medication information in patient registries. Arthritis Rheum, 2007. **57**(2): p. 234-39.

51. Merkin, S.S., et al., *Agreement of self-reported comorbid conditions with medical and physician reports varied by disease among end-stage renal disease patients.* J Clin Epidemiol, 2007. **60**(6): p. 634-42.

52. Boissonnault, W.G. and M.B. Badke, *Collecting health history information: the accuracy of a patient self-administered questionnaire in an orthopedic outpatient setting.* Phys Ther, 2005. **85**(6): p. 531-43.

53. Iversen, L., et al., *Do people self-reporting information about chronic respiratory disease have corroborative evidence in their general practice medical records? A study of intermethod reliability.* Prim Care Resp J, 2007. **16**(3): p. 162-68.

54. Corser, W., et al., *Concordance between comorbidity data from patient self-report interviews and medical record documentation.* BMC Health Services Research, 2008. **8**(1): p. 85.

55. Baptiste-Roberts, K., B. Barone, and T. Gary, *Risk factors for type 2 diabetes among women with gestational diabetes: a systematic review.* Am J Med, 2009. **122**(3): p. 207-214.

56. Baraban, E., L. McCoy, and P. Simon, *Increasing prevalence of gestational diabetes and pregnancy-related hypertension in Los Angeles County, California, 1991–2003.* Prev Chronic Dis, 2008. **5**(3).

57. Gagne, J.J., et al., *Prescription drug use during pregnancy: a population-based study in Regione Emilia-Romagna, Italy.* Eur J Clin Pharmacol, 2008. **64**(11): p. 1125-1132.

58. Malm, H., et al., *Prescription drugs during pregnancy and lactation--a Finnish register-based study.* Eur J Clin Pharmacol, 2003. **59**(2): p. 127-133.

59. Bracken, M. and T. Holford, *Exposure to prescribed drugs in pregnancy and association with congenital malformations.* Obstet Gynecol, 1981. **58**: p. 336-344.

60. Leikauf, J. and A.D. Federman, *Comparisons of Self-Reported and Chart-Identified Chronic Diseases in Inner-City Seniors.* J Am Geriatr Soc, 2009. **57**(7): p. 1219-25.

61. Brownson, R.C., et al., *Reliability of information on chronic disease risk factors collected in the Missouri Behavioral Risk Factor Surveillance System.* Epidemiology, 1994. **5**(5): p. 545-49.

62. Martin, L.M., et al., *Validation of self-reported chronic conditions and health services in a managed care population.* Am J Prev Med, 2000. **18**(3): p. 215-18.

63. Miller, D.R., et al., *Patients' Self-report of Diseases in the Medicare Health Outcomes Survey Based on Comparisons With Linked Survey and Medical Data From the Veterans Health Administration.* J Ambul Care Manage, 2008. **31**(2): p. 161-77.

64. Young, T.K., L.L. Roos, and D.E. Gelskey, *Estimating the burden of disease: comparing administrative data and self-reports.* Med Care, 1997. **35**(9): p. 932-47.

65. Kwon, A., et al., *Antidepressant use: concordance between self-report and claims records.* Med Care, 2003. **41**(3): p. 368-74.

66. Sjahid, S.I., P.D. van der Linden, and B.H.C. Stricker, *Agreement between the pharmacy medication history and patient interview for cardiovascular drugs: the Rotterdam elderly study.* Br J Clin Pharmacol, 1998. **45**(6): p. 591-95.

67. Brown, D.W., R.F. Anda, and V.J. Felitti, *Self-reported information and pharmacy claims were comparable for lipid-lowering medication exposure.* J Clin Epidemiol, 2007. **60**: p. 525-29.

68. Caskie, G.I.L. and S.L. Willis, *Congruence of self-reported medications with pharmacy prescription records in low-income older adults.* The Gerontologist, 2004. **44**(2): p. 176-85.

69. Haukka, J., et al., *High concordance between self-reported medication and official prescription database information.* Eur J Clin Pharmacol, 2007. **63**(11): p. 1069-74.

70. Nielsen, M.W., et al., *Agreement between self-reported data on medicine use and prescription records vary according to method of analysis and therapeutic group.* J Clin Epidemiol, 2008. **61**(9): p. 919-24.

71. Kropp, S., et al., *Good agreement between physician and self-reported hormone therapy data in a case–control study.* J Clin Epidemiol, 2007. **60**: p. 1280-87.

72. Løkkegaard, E.L., et al., *The validity of self-reported use of hormone replacement therapy among Danish nurses.* Acta Obstet Gynecol Scand, 2004. **83**(5): p. 476-81.

73. Curtis, J., et al., *Agreement and validity of pharmacy data versus self-report for use of osteoporosis medications among chronic glucocorticoid users.* Pharmacoepidemiology and Drug Safety, 2006. **15**(10): p. 710-18.

74. Abdolrasulnia, M., et al., *Agreement between medical record documentation and patient-reported use of nonsteroidal antiinflammatory drugs.* Am J Health Syst Pharm, 2006. **63**(8): p. 744-47.

75. West, S.L., et al., *Demographics, health behaviors, and past drug use as predictors of recall accuracy for previous prescription medication use.* J Clin Epidemiol, 1997. **50**(8): p. 975-80.

76. Kelly, J.P., et al., *Reliability of personal interview data in a hospital-based case-control study.* Am J Epidemiol, 1990. **131**(1): p. 79-90.

77. Britton, G.R.A., et al., *Comparison of self-reported smoking and urinary cotinine levels in a rural population.* JOGNN, 2003. **33**: p. 306-11.

78. Rice, F., et al., *Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics.* Early Human Development, 2007. **83**(8): p. 497-504.

79. Hessol, N.A., B. Missett, and E. Fuentes-Afflick, *Lower agreement on behavioral factors than on medical conditions in self-reported data among pregnant Latina women.* Arch Med Res, 2004. **35**(3): p. 241-45.

80. Fox, N.L., et al., *The reliability of self-reports of smoking and alcohol consumption by pregnant women.* Addict Behav, 1989. **14**: p. 187-95.

81. Jacobson, S.W., et al., *Maternal recall of alcohol, cocaine, and marijuana use during pregnancy.* Neurotoxicol Teratol, 1991. **13**(5): p. 535-40.

82. Olson, J.E., et al., *Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group.* Am J Epidemiol, 1997. **145**(1): p. 58-67.

83. Bosetti, C., et al., *Reliability of data on medical conditions, menstrual and reproductive history provided by hospital controls.* J Clin Epidemiol, 2001. **54**: p. 902-06.

84. Vilian, A., et al., *Agreement between administrative databases and medical charts for pregnancy-related variables among asthmatic women.* Pharmacoepidemiol Drug Saf, 2008. **17**: p. 345-53.

85. Tilley, B.C., et al., *A comparison of pregnancy history recall and medical records: implications for retrospective studies.* Am J Epidemiol, 1985. **121**(2): p. 269-81.

86. Buka, S.L., et al., *Maternal recall of pregnancy history: accuracy and bias in schizophrenia research.* Schizophrenia Bulltin, 2000. **26**(2): p. 335-50.

87. Hasan, R., et al., *Accuracy of reporting bleeding during pregnancy.* Paediatric and Perinatal Epidemiology, 2010. **24**: p. 31-34.

88. De Jong, P.C.M., et al., *Validation of a questionnaire on medical drug use during pregnancy.* Am J Epidemiol, 1991. **134**(9): p. 998-1002.

89. Mitchell, A.A., L.B. Cottler, and S. Shapiro, *Effect of questionnaire design on recall of drug exposure in pregnancy.* Am J Epidemiol, 1986. **123**(4): p. 670-76.

90. Klungel, O.H., et al., *Influence of question structure on the recall of self-reported drug use.* J Clin Epidemiol, 2000. **53**(3): p. 273-77.

91. http://www.cdc.gov/nchswww/fastats/fastats.htm, C.f.D.C.a.P.N.C.f.H.S.F.R.f.t.W.W.W. 2009 [cited 2009 September 9].

92. Perrin, E.C., et al., *"Issues Involved in the Definition and Classification of Chronic Health Conditions,"*. Pediatrics., 1993. **91**: p. 787-93.

93. *Drug Facts and Comparisons 4.0* Vol. 1. 2008: Wolters Kluwer.

94. Landis, J.R. and G.G. Koch, *The measurement of observer agreement for categorical data.* Biometrics, 1977. **33**(1): p. 159-174.

95. Feinstein, A.R. and D.V. Cicchetti, *High agreement but low Kappa: I. the problems of two paradoxes* 1.* J Clin Epidemiol, 1990. **43**(6): p. 543-49.

96. Byrt, T., J. Bishop, and J.B. Carlin, *Bias, prevalence and kappa.* J Clin Epidemiol, 1993. **46**(5): p. 423-29.

97. Sim, J. and C.C. Wright, *The kappa statistic in reliability studies: use, interpretation, and sample size requirements.* Phys Ther, 2005. **85**(3): p. 257-68.

98. Altman, D.G. and J.M. Bland, *Diagnostic tests. 1: Sensitivity and specificity.* BMJ, 1994. **308**(6943): p. 1552.

99. Hosmer, D.W. and S. Lemeshow. *Applied Logistic Regression*. 1st edition ed. 1989, John Wiley & Sons, Inc.: New York.

100. *Department of Health and Human Services Public Health Service. Surgeon General's advisory on alcohol and pregnancy.* FDA Drug Bulletin, 1981. **11**(2): p. 9-10.

101. Merlo, J., et al., *Self-administered questionnaire compared with a personal diary for assessment of current use of hormone therapy: An analysis of 16, 060 women.* Am J Epidemiol, 2000. **152**(8): p. 788-92.

102. Sandini, L., et al., *Agreement of self-reported estrogen use with prescription data: an analysis of women from the Kuopio Osteoporosis Risk Factor and Prevention Study (OSTPRE).* Menopause, 2008. **15**(2): p. 282-89.

103. Hessol, N.A., et al., *Accuracy of self-reports of acquired immunodeficiency syndrome and acquired immunodeficiency syndrome-related conditions in women.* Am J Epidemiol, 2001. **153**(11): p. 1128-33.

104. Hersh, W.R., *Adding value to the electronic health record through secondary use of data for quality assurance, research, and surveillance.* Am J Manag Care, 2007. **81**: p. 126-28.

105. Pakhomov, S., et al., *Agreement between patient-reported symptoms and their documentation in the medical record.* Am J Manag Care, 2008. **14**(8): p. 530-39.

106. Wang, P.S., et al., *How well do patients report noncompliance with antihypertensive medications?: a comparison of self-report versus filled prescriptions.* Pharmacoepidemiology and Drug Safety, 2004. **13**(1): p. 710-18.

107. Lester, B.M., et al., *The Maternal Lifestyle Study: drug use by meconium toxicology and maternal self-report.* Pediatrics, 2001. **107**(2): p. 309-17.