

Finding Understudied Disorders Potentially Associated with Maternal Morbidity and Mortality

Laritza M. Rodriguez, MD, PhD¹ Dina Demner Fushman, MD, PhD¹

¹Lister Hill National Center for Biomedical Communications, National Library of Medicine, National Institutes of Health, Bethesda, Maryland

Address for correspondence Laritza Rodriguez, MD, PhD, Lister Hill National Center for Biomedical Communications, National Library of Medicine, 9000 Rockville Pike, Bethesda, MD 20892 (e-mail: laritza.rodriguez@nih.gov).

Am J Perinatol Rep 2019;9:e36–e43.

Abstract

Objective Clinical research literature focuses primarily on the most common causes of maternal morbidity and mortality (MMM). We explore sections of the discharge summaries of pregnant or postpartum women admitted to an intensive care unit (ICU) to identify associated disorders and mine the literature to identify knowledge gaps in clinical research.

Methods Data for the study were discharge summaries in the MIMIC (Medical Information Mart for Intensive Care) database. We extracted a control cohort to study if there is a difference in comorbidities between pregnant and not pregnant patients with similar reasons for admission. We identified comorbidities of the Unified Medical Language System (UMLS) semantic types disease or syndrome, Mental or behavioral dysfunction, and injury, or poisoning. We used Entrez programming utilities (E-utilities) to query PubMed®.

Results We identified 246 pregnant and postpartum patients. A control group of 587 not pregnancy related admissions matched on age and admit diagnosis. We found overlap of 24.3% discharge diagnoses between the two groups, and 7.5% of the codes exclusively in the pregnancy group. We identified 33 disease mentions not included in the most common reported causes of MMM.

Conclusion Our results demonstrate that clinical text provides additional comorbidities associated with maternal complications that need further clinical research.

Keywords

- ▶ maternal morbidity
- ▶ maternal mortality
- ▶ puerperium
- ▶ preventing maternal deaths

The center for disease control (CDC) defined a pregnancy related death as the death of a woman while pregnant or within 1 year of pregnancy termination regardless of the outcome, the duration or site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes.¹ Maternal comorbidities are medical pregnancy complications caused by the pregnancy itself, for example, pregnancy-induced hypertension or underlying diseases aggravated or concurrent with the pregnancy, such as chronic hyperten-

sion. Maternal morbidity and mortality (MMM) and birth outcomes are population health measures that reflect the overall health condition of a population. In general, populations with higher standards of care and quality of life should have lower ratios of MMM. However, between 1990 and 2013 in the U.S.A., the expected 12 maternal deaths per 100,000 live births more than doubled, causing the country to be at the highest ratios of maternal deaths among other high income countries.² Recent reports estimate that each year 1,200 women in the U.S.A. die from pregnancy-related

received
January 15, 2019
accepted after revision
January 26, 2019

DOI <https://doi.org/10.1055/s-0039-1683363>.
ISSN 2157-6998.

Copyright © 2019 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New York, NY 10001, USA.
Tel: +1(212) 584-4662.

License terms



complications, and approximately 60,000 are suffering from near-fatal complications. In December 2018, the bill introduced by the United States Congress “Preventing Maternal Deaths Act of 2018” was signed into law.³ The focus of this law is to increase federal support of States, Indian tribes, and tribal organizations to record and report pregnancy associated deaths, and to implement the recommendations from the review committees to prevent maternal deaths.

Due to the low incidence of maternal death and medical complications in the pregnant population, it has always been difficult to obtain enough data for meaningful analysis. To compensate for the low incidence, the analysis has been expanded to include severe acute maternal morbidity (SAMM), also known as near-miss morbidity and mortality.⁴ According to the World Health Organization (WHO), near-miss obstetrical events are those in which a pregnant or recently postpartum woman survives a life-threatening event either by chance or because of high quality medical care.⁵ The identification of such cases and the inclusion criteria are not always straightforward; the complexity of the cases and differences in management criteria and availability of resources across institutions and in different countries could create confusion in the inclusion criteria. Some authors define MMM based on organ system-based criteria or end-organ damage, others, based on disease severity criteria, and others yet focus on management based criteria, and intensive care unit (ICU) admissions.⁶ Reports in the literature have aimed to define specific inclusion criteria but the overall conclusion is that there is no single best approach; the authors should aim to include the highest number of patients for analysis.⁷

The most common globally reported causes of maternal morbidity and mortality are peripartum hemorrhage, eclampsia, and complicated preeclampsia, obstructed birth, and sepsis.⁸ The latest official report on maternal mortality in the U.S.A. was published in 2007. Since then, several studies have reported the increase in maternal mortality in the U.S.A.⁹ To improve detection, in 2003, the CDC added a pregnancy specific section to the death certificate including specific questions on time of pregnancy in reference to the date of death. There was considerable delay in the implementation of the questions in several States, and other States added questions that do not comply with the CDC standard. Consequently, there is no clear knowledge to date about the trends on maternal mortality in the country.⁹ Moreover, studies focusing on the near-miss cases in the U.S.A. are sparse, and the analyses of comorbidities and pregnancy-associated diseases which complicate the outcomes, are lacking in the literature.

A PubMed® search for “((near miss AND pregnancy)) AND (“2012”[Date - Publication]: “2017” [Date - Publication])” returns 193 studies in the last 5 full calendar years with only one based on USA population, with data recorded from 2005–2010. Two large studies conducted in Texas and Maryland and focusing on intensive care admissions of pregnant women, report on pregnancy related diseases with little mention of coexisting diseases complicating the pregnancy.^{10,11} The two studies emphasize that the admission of a pregnant or postpartum woman alone is indicative of

SAMM, and the focus on admissions to the ICU includes close to 100% of patients with pregnancy complications. Most importantly, there are no decision support systems for early identification of pregnant women with comorbidities potentially leading to near-miss cases. Such systems are needed for adequate triaging of pregnant women.

Compounded by hormonal, immunological, hemodynamic, and physical changes in the human body, pregnancy is a physiologic state that alters the normal physiology and has the potential for overloading organisms that might be already functioning at maximum capacity. To make matters worse, there are weight, vascular and intravascular changes that alter the response to established effective treatments. In addition, medications that are effective and perfectly safe in a nonpregnant woman might be contraindicated in pregnancy due to potential effects on the embryo and fetus. The majority of pregnant women attend prenatal clinics where the attending personnel may lack specialized training in high-risk pregnancies. It is important to provide tools to these practitioners for identification and management of highly complex medical and obstetrical cases. The improvement in medical care and changes in women’s lifestyles have increased not only the average childbearing age but also the number of pregnant women with chronic medical conditions. These women are more likely to have complicated pregnancies and deliveries, and to require intensive care unit admissions. The obstetrics and maternal–fetal literature are rich in research on the most frequent causes of maternal morbidity and peripartum complications, such as preeclampsia, postpartum hemorrhage, premature rupture of membranes, and diabetes. Studies on less common comorbidities are lacking.

Our previous research demonstrates that obstetrical data are notoriously sparse on comorbidities.¹² Previous reports have identified knowledge gaps in the scientific literature. In a systematic review, Gil-González et al reported the lack of published information about cultural and political determinants of maternal mortality, and the lack of reports from underdeveloped countries.¹³ Knowledge gaps are “inverse scientific interest pattern between the burden of the causes of maternal mortality and scientific interest in these determinants.¹²” To the best of our knowledge, there are no studies on knowledge gaps in the literature based on clinical comorbidities. We previously demonstrated that it is possible to identify unknown associations in the past medical history for Lyme disease.¹⁴ We aim to explore the generalizability of the methodology we will be proposing, using the same approach to find diseases associated with maternal complications that need further research and are underrepresented in the literature. We explore sections of the discharge summaries of pregnant or postpartum women admitted to an ICU to identify associated diseases and disorders and then mine the literature to identify knowledge gaps in clinical research.

Methods

The data for the study consisted of discharge summaries from pregnant or peripartum patients from MIMIC (medical

information mart for intensive care) III and MIMIC II databases.¹⁵ The patients were identified using Essie, the concept-based search engine,¹⁶ searching the clinical notes for “pregnancy or pregnant.” We used broad search terms to ensure the retrieval of all patients with pregnancy-related events. Our search returned 3,679 unique patients. Manual review of the clinical notes, radiology reports, and nursing notes resulted in 246 (6.7%) pregnant or postpartum patients or women who had ectopic pregnancies or spontaneous abortions. The remaining 3,433 (93.3%) records were from newborns, male patients, mention of a pregnant relative, suspected pregnancies, or past medical history of pregnancies. We extracted the clinical notes for 178 patients from MIMIC III, and clinical notes for 68 patients from MIMIC II (total patients 246). According to the documentation, MIMIC III is the most up to date version of the database and inclusive of all the data in MIMIC II. However, we were unable to find all the clinical notes for all the relevant patients in the latest version. Therefore, we used both databases, as well as the most recent version of each document. We extracted the demographic and procedure data from the structured tables in the database. The ICD-9 (international classification of diseases, 9th revision) discharge diagnoses were mapped into general clinical categories based on the health care cost and utilization project (HCUP) clinical classification software (CCS).¹⁷ The CCS is a tool for clustering patient diagnoses and

procedures into clinically meaningful categories to facilitate the analysis of associated illnesses. CCS collapses diagnosis and procedure codes from the ICD-9 clinical modification (ICD-9-CM) which contains more than 14,000 diagnosis codes and 3,900 procedure codes. The single level diagnosis clinical classification aggregates ICD-9 codes of illnesses and conditions into 285 mutually exclusive categories. For example, ICD-9 codes 4011–4019 (multiple level diseases) are classified into a single level clinical category: Essential hypertension, clinical category 98. We used the single level disease categories to identify the ICD-9 codes relevant to pregnancy. The single level disease categories (codes) used are shown in ►Table 1.

We extracted a control cohort to study whether there is a difference in comorbidities between pregnancy associated ICU admissions and nonpregnancy related admissions. The control cohort was extracted from the control group of all female patients in the database between the ages 16 and 52 years (8,134 admissions and 4,122 distinct patients). We used the “diagnosis” column of the “admissions table” in MIMIC III. We matched the 246 study patients to the controls on the admission diagnosis and age. On average, we found two controls for each study subject. For those patients in the pregnancy group where the admission diagnosis was only relevant to pregnancy, we manually extracted the “ICU reason for admission” from the discharge summaries to

Table 1 Disease categories for mapping discharge diagnoses

General clinical categories	Health care cost and utilization project (HCUP) code
Spontaneous abortion	177
Induced abortion	178
Postabortion complications	179
Ectopic pregnancy	180
Other complications of pregnancy	181
Hemorrhage during pregnancy, placenta abruption, placenta previa	182
Hypertension complicating pregnancy, child birth, and the puerperium	183
Early or threatened labor	184
Prolonged pregnancy	185
Diabetes or abnormal glucose tolerance complicating pregnancy, childbirth, or the puerperium	186
Malposition, malpresentation	187
Fetopelvic disproportion, obstruction	188
Previous C-section	189
Fetal distress and abnormal forces of labor	190
Polyhydramnios and other problems of amniotic cavity	191
Umbilical cord complication	192
OB-related trauma to perineum and vulva	193
Forceps delivery	194
Other complications of birth; puerperium affecting management of mother	195
Other pregnancy and delivery including normal	196

match these disorders to the diagnoses in the controls. For example, in a study group, the patient reason for admission was 39 weeks' induction. This patient had an ICU reason for admission—uncontrolled hypertension, which we matched to hypertension emergency in the control group. We used the ICD-9 discharge codes grouped at the first CCS level disease category to compare comorbidity diagnoses in the two groups.

To see whether the clinical notes provide additional important comorbidities, we annotated clinical notes. Although we intended to annotate all categories of clinical notes, we found that the discharge summaries were the most inclusive of the medical history and contained all information needed for this study. After discarding duplicates, we were left with 272 discharge summaries from 246 patients in the study group, including readmissions.

We split the discharge summaries into sections using a rule-based system.¹⁸ We used text from the history or present illness, hospital course, and discharge diagnosis sections. Using the natural language tool kit (NLTK),¹⁹ sentence segmenter to extract sentences from these sections, we tagged each sentence with its subject, admission, and discharge summary unique identification numbers, as well as section name, section number, and sentence number. We manually annotated the above sentences using BRAT annotation tool.²⁰ We designed the annotation schema to extract demographics, medications, diseases and disorders, past medical history, procedures, discharge destinations, and discharge conditions. For the purpose of the present study, we annotated all concepts potentially related to SAMM: diagnosis on admission, cause of death, chief complain, discharge diagnoses, massive transfusion, non-pregnancy related surgical procedures, postpartum complications, problem list, reason for ICU admission, and substance abuse.

We then intersected the diseases extracted from the records of the study and control groups and excluded the common diseases from further consideration: even if these disease are understudied, they are understudied for female ICU patients in general.

The next step was to explore the amount of scientific research interest in the diseases or disorders in ICU admitted patients classified as near-miss obstetrical cases and maternal mortality that were retrieved by our methods. We focused on terms in the unified medical language system (UMLS)²¹ with semantic types disease or syndrome, mental or behavioral dysfunction, injury, or poisoning. Terms with other semantic types, for example, sign or symptom are too general and not precise enough for literature searches, such as nausea, vomiting, fever, and pain.

We then used Entrez programming utilities (E-Utilities) from the National Center for Biotechnology Information (NCBI)²² to query PubMed,²³ limiting the searches to the past 5 years of published literature. The queries were constructed using the concepts verified by manual annotation as noted above, and pregnancy. We used the *EFetch* function to download the full records. For example, a query was constructed from the following discharge summary text: "A diagnosis of abdominal compartment syndrome was made and the patient had taken to the operating room for an

emergent decompressive laparotomy." UMLS preferred term "abdominal compartment syndrome"; CUI (Concept Unique Identifier): C1142110 disease or syndrome. E-utilities search: "abdominal compartment syndrome" (all fields) and ("pregnancy" [MeSH Terms] or "pregnancy" [all fields]), and ("2013/01/01" [PDat]: "2013/12/31" [PDat])."

Results

We identified 246 pregnant and postpartum patients between 16 and 52 years old, with an average age of 33. The patients were White (45%), African American (23%), and other or unknown (32%). The length of stay (LOS) ranged from less than 1 day to 92 days, with an average LOS of 9.7 days. Patients had private insurance (59%), Medicaid (30%), and other insurance (11%). About two-third (66%) of the patients were married, and 34% were single or divorced. The discharge diagnoses based on ICD-9 codes and clinical categories were as follows: peripartum hemorrhage or anemia (19.93%); sepsis or infection (14.55%); cardiovascular disease or disorder (6%); chronic or pregnancy induced hypertension (5.89%); renal or urinary disease or disorder (5.61%); trauma, for example, falls or motor vehicle accident (3.28%); diabetes (3.00%); medical or surgical procedure complication (2.94%; ► **Table 2**). The ICD-9 codes also include descriptions, such as postterm pregnancy, delivered, which are confirmation of a pregnancy-related admission. However, we do not include the codes in ► **Table 2**, as we only report on comorbidities and we have confirmation of pregnancy related admission from the manual chart review. The counts of the ICD-9 codes indicate the number of times the code is in the database for the patient cohort and not the number of patients with the condition. One patient can have several ICD-9 codes in the same clinical category, such as hemorrhage or anemia coded as hemorrhage or anemia complicating a procedure (998.11), and/or delayed and secondary postpartum hemorrhage or anemia, postpartum condition or complication (666.24). Likewise, one patient can have several diagnoses that cascade during the hospital stay, for example, pregnant patient admitted for preeclampsia could have had severe postpartum hemorrhage that required massive transfusion and caused postpartum coagulation defects, with the patient then developing sudden adult respiratory failure (SARS), and becoming septic. The average number of ICD-9 codes per patient is 10 with a maximum of 69 and one patient had only one code (eclampsia, antepartum condition, or complication [642.63]).

There were 11 deceased patients, 5 of which died within 1 year of the pregnancy termination and are the only ones classified as maternal mortality. The underlying causes of death for these patients were peripartum cardiomyopathy,² suicide and self-inflicted injury by hanging,¹ diabetes with ketoacidosis, type I,¹ and malignant neoplasm of head of pancreas.¹ Maternal death rate is 1:49 in the study cohort, ratio of deceased patients to those admitted to the ICU with pregnancy-related events in 1 year in the study cohort. This number is not a population-based measure and only indicates the mortality rate in our study cohort.

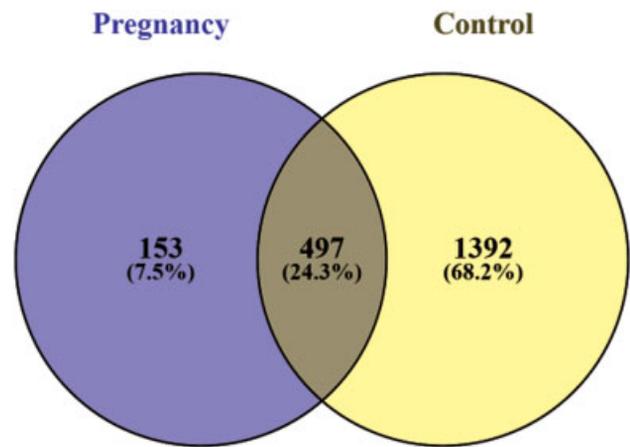
Table 2 Clinical classification categories for ICD-9 codes not including pregnancy-related codes

Clinical Categories	n	%
Hemorrhage or anemia	359	19.93
Sepsis or Infection	262	14.55
Cardiovascular disease or disorder	108	6.00
Renal disease or urinary condition	101	5.61
Trauma	59	3.28
Diabetes	54	3.00
Procedure complication	53	2.94
Sars	52	2.89
Mental depression or suicide	46	2.55
Cerebrovascular or neurological disorders	39	2.17
Thyroid	32	1.78
Asthma	29	1.61
Malignancy	24	1.33
Cardiomyopathy	24	1.33
Pelvic organ injury or laceration	23	1.28
Substance abuse	17	0.94
Shock	16	0.89
Immunological disease	12	0.67
ADE	9	0.50
Hypovolemia	7	0.39
Organ transplant status	5	0.28
Renal disease or urinary condition	2	0.11

Abbreviations: ADE, adverse drug event; ICD-9, international classification of diseases, 9th revision.

In the control cohort, there were 4,122 patients with a matching age range. Matching on admission diagnosis reduced this group to 587 patients, with an average 10-day LOS, maximum stay of 141 and minimum of 0 days, and 32 deceased patients. The study group mapped to 190 CCS single level diagnoses groups, and the control group to 238 single level diagnoses groups. The study group had 842 distinct ICD-9 codes (620 not pregnancy-related distinct codes) versus 1,808 distinct codes in the control group. The overlap of the ICD-9 distinct codes was 24.3%, with 7.5% of the codes exclusively in the pregnancy group and 68.2% in the control group (► Fig. 1).

The top 10 single level diagnoses groups from the single level disease categories in the study group (not including pregnancy related codes) were *fluid and electrolyte disorders*, *deficiency and other anemia*, acute posthemorrhagic anemia, respiratory failure, insufficiency, arrest (adult), *septicemia (except in labor)*, coagulation and hemorrhagic disorders, *complications of surgical procedures or medical care*, congestive heart failure, nonhypertensive, urinary tract infections, and asthma. The top 10 single level diagnosis groups in the control group were diabetes mellitus with complications, *fluid and electrolyte disorders*, *deficiency and other anemia*,

**Fig. 1** Results of overlap of ICD-9 discharge diagnosis from the pregnant and not pregnant groups. ICD-9, international classification of diseases, 9th revision.

septicemia (except in labor), residual codes, unclassified, *complications of surgical procedures or medical care*, other nervous system disorders, mood disorders, complication of device, implant or graft, chronic kidney disease. Only 4 of the 10 top groups overlap (indicated above in italics) between the study and control groups.

The manual review of the clinical notes resulted in 33 disease mentions not included in the most common reported causes of MMM. From these, 61% were included in the ICD-9 diagnosis codes, 39% were extracted by manual review only.

The results from the E-utilities searches in PubMed for the past 5 years for the retrieved terms and pregnancy are presented in ► Table 3.

Discussion and Conclusions

Our results demonstrate that clinical text provides additional comorbidities associated with maternal complications that need further clinical research, although 61% of the diseases mentioned in the notes were also found in the structured data, there were still 39% of understudied diseases identified by manual review only. Together, the notes and structured data help find understudied disorders potentially associated with pregnancy and delivery. With the use of automatic methods enhanced with manual review, we extracted a cohort of pregnant and postpartum patients. We identified pregnancy-associated diseases not reported as the most common causes of MMM in the cohort, and we used the results with E-utilities to explore the clinical literature. The results indicate that the literature predominantly reports on clinical research of the most common causes of MMM but is lacking in other areas where the clinical incidence of the associated diseases is not as high but still important enough to cause serious maternal complications and even death.

Our results also demonstrate that it is possible to make use of clinical records to find disorders underrepresented in the literature and to present the results of clinical research on patients admitted to the ICU during pregnancy or postpartum.

Table 3 Results by year of PubMed® for the retrieved disease terms associated to pregnancy or postpartum intensive care unit admissions, not including the most common causes of MMM reported in the literature

Retrieved terms associated with pregnancy ICU Admissions	Year					
	2013	2014	2015	2016	2017	2018
Abdominal compartment syndrome	3	12	7	4	5	
Acute renal failure	59	40	74	74	62	8
Asthma		222	231	278	232	32
Brain injury	110	151	136	128	75	4
Cardiomyopathy	108	154	143	147	98	5
Deep venous thrombosis	121	149	142	144	108	6
Diabetic ketoacidosis	10	15	6	11	13	
Drug abuse		23	417	345	206	6
Drug overdose ^a	10	9	9	13	5	
Duodenal ulcer ^a	1	1	1	1		
Endomyometritis	45	68	62	55	48	3
G6PD deficiency ^a	5	3	5	6	6	
Gout	4	2	4	6	6	
HIV				297	657	43
Intraventricular hemorrhage	56	48	65	72	50	5
Liver rupture	10	15	15	19	14	1
Massive transfusion ^a	1	4	3	4	3	
Mental disease			316	479	198	6
Myasthenia	15	6	13	11	10	4
Ovarian hyperstimulation syndrome ^a	91	105	95	93	93	8
Pseudocyst of pancreas ^a				1		
Pericardial effusion ^a	9	14	12	11	6	
Pleural Effusion	14	17	19	11	13	
Porphyria ^a	5	4	1	2	1	
Posterior reversible leukoencephalopathy ^a syndrome	13	14	14	7	8	
Pulmonary embolism	76	82	103	79	65	7
SARS	64	78	87	78	48	
Severe perineal laceration	9	23	13	15	5	
Suicide	45	39	52	65	37	4
Thyrotoxicosis ^a	13	24	22	16	9	1
Trauma accident	42	31	31	29	13	
Venous sinus thrombosis ^a	12	13	16	10	14	
Vertebral artery aneurysm ^a	1	4	2	2		
Total	952	1,370	2,116	2,513	2,108	143

Abbreviations: G6PD, glucose-6-phosphate-dehydrogenase deficiency; HIV, human immunodeficiency virus; ICU, intensive care unit; MMM, maternal morbidity and mortality; SARS, sudden adult respiratory failure;

^aIndicates diseases extracted only by manual review of clinical notes.

We explored a database of ICU patients with combined manual and automatic methods to extract information on pregnant and postpartum patients. We verified pregnancy and postpartum status using both coded data (ICD-9 discharge diagnosis), automatic chart retrieval with search terms, and manual review of clinical notes. We found that

pregnancy mentions in clinical notes or ICD-9 codes alone are not reliable for patient classification. As reported in the literature and verified by our results, the incidence of maternal complications that require admission to the ICU is very low. In spite of the big size of the ICU database, we only had 246 verified patients that fitted the criteria of our study.

The main causes of MMM and near-miss cases reported in the literature and confirmed in our results are postpartum hemorrhage, complications of preeclampsia, and diabetes but there are other associated diseases that require additional clinical study specifically in the pregnancy setting, such as cardiomyopathy, deep venous thrombosis, intraventricular hemorrhage, vertebral artery aneurysm, and mental disease among others.

The comparison analysis of the ICD-9 codes and single level disease groups between the study group and a control cohort show an overlap of ICD-9 distinct codes (24.3%). These results suggest that by generalizing disease diagnoses, the finer granularity of disease specific groups is lost.

Our study also confirms that the lack of structure and consistency in the document sections create challenges. The documents have the same sections repeated or spread within the documents with no systematic order. Some discharge summaries include transcripts from other types of reports (radiology reports embedded in the discharge summary with no clear section break). Not all the reports include all sections and often the sections are in different order. It is not uncommon to find snippets of past medical history mentioned in the history of present illness when the past history is relevant to the present condition. For example, history of recent tobacco or drug use in a pregnant woman with shortness of breath. Other known problems of free text are frequent misspellings, use of nonconventional abbreviations, mentions of bacteria or abnormal cells to indicate a disorder, for example, schistocyte, to indicate the severity of an intravascular coagulation state, or abnormal laboratory, or test results “CD-4 of 1360” with no indication of the name of the disorder that these are associate with.

The nature of the ICU admissions also creates complexity, for example, postpartum women admitted to the ICU are often transferred after a cesarean section, or after a massive hemorrhage in labor and delivery that required a hysterectomy and massive transfusion. In these cases, the history of present illness is “postcesarean hysterectomy,” with UMLS semantic type therapeutic or preventive procedure, not a disorder nor a diagnosis but a condition that prompted the patient admission to the ICU and therefore, that is, the patient status on admission.

Coded clinical data are desirable for large population analysis but it is necessary to resort to manual and natural language processing methods to extract hidden knowledge in the unstructured data. Manual processing of clinical text is time consuming and does not allow to process thousands of documents to identify a small cohort of patients with rare but potentially deadly pregnancy complications. Although we used database queries and literature search tools to assist our study, more automation and wider access to the lists of diseases co-occurring with peripartum are needed to identify the diseases associated with pregnancy those are underrepresented. In this study, we intentionally omitted the use of automated extraction of co-occurring conditions from clinical text, as our previous study showed that approximately 20% of the disorders will be missed, and we wanted to establish more reliably the actual contribution of clinical text.

Limitations and Future Work

Our analysis is limited due to the relatively small size of the sample, and the low incidence of diseases in the pregnant population. The data source used is an administrative database with deidentified records which compounds the limitations of the analysis. We hope in the future to improve automatic methods for free text document segmentation and extraction of disorders from clinical text, so as to fully automate the clinical record-based approach to finding understudied disorders associated with pregnancy.

Conflict of Interest

None.

Acknowledgments

This work was supported by the Intramural Research Program of the National Library of Medicine, National Institutes of Health.

References

- 1 Pregnancy mortality surveillance system. Centers for disease control and prevention. Available from: <https://www.cdc.gov/reproductivehealth/maternalinfanthealth/pmss.html>. Accessed February 11, 2019
- 2 Main EK, Menard MK. Maternal mortality: time for national action. *Obstet Gynecol* 2013;122(04):735–736
- 3 Preventing maternal deaths act of 2018. One hundred fifteenth congress of the United States of America. 2019. Available from: <https://www.congress.gov/bill/115th-congress/house-bill/1318/text>. Accessed February 11, 2019
- 4 Mhyre JM, Bateman BT, Leffert LR. Influence of patient comorbidities on the risk of near-miss maternal morbidity or mortality. *Anesthesiology* 2011;115(05):963–972
- 5 Evaluating the quality of care for severe pregnancy complications. World Health Organization; 2011. Available from: http://apps.who.int/iris/bitstream/10665/44692/1/9789241502221_eng.pdf. Accessed February 11, 2019
- 6 Say L, Souza JP, Pattinson RC; WHO working group on Maternal Mortality and Morbidity classifications. Maternal near miss—towards a standard tool for monitoring quality of maternal health care. *Best Pract Res Clin Obstet Gynaecol* 2009;23(03):287–296
- 7 Clark SL, Meyers JA, Frye DR, McManus K, Perlin JB. A systematic approach to the identification and classification of near-miss events on labor and delivery in a large, national health care system. *Am J Obstet Gynecol* 2012;207(06):441–445
- 8 Bale J, Stoll B, Lucas A. Improving Birth Outcomes: Meeting the Challenge in the Developing World. Committee on Improving Birth Outcomes Institute of Medicine (US). 1st ed. Washington DC, WA: National Academic Press; 2003
- 9 MacDorman MF, Declercq E, Cabral H, Morton C. Recent increases in the U.S. maternal mortality rate: disentangling trends from measurement issues. *Obstet Gynecol* 2016;128(03):447–455
- 10 Oud L. Epidemiology of pregnancy-associated ICU utilization in Texas: 2001 - 2010. *J Clin Med Res* 2017;9(02):143–153
- 11 Wanderer JP, Leffert LR, Mhyre JM, Kuklina EV, Callaghan WM, Bateman BT. Epidemiology of obstetric-related ICU admissions in Maryland: 1999–2008*. *Crit Care Med* 2013;41(08):1844–1852
- 12 Taft LM, Evans RS, Shyu CR, et al. Countering imbalanced datasets to improve adverse drug event predictive models in labor and delivery. *J Biomed Inform* 2009;42(02):356–364
- 13 Gil-González D, Carrasco-Portiño M, Ruiz MT. Knowledge gaps in scientific literature on maternal mortality: a systematic review. *Bull World Health Organ* 2006;84(11):903–909

- 14 Rodriguez L, Demner Fushman D. Extracting diseases co-occurring with past medical history of Lyme disease. AMIA Annu Symp Proc 2018. Available at: <https://informaticssummit2018.zerista.com/event/member/470616>. Accessed February 21, 2019
- 15 Saeed M, Villarroel M, Reisner AT, et al. Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database. Crit Care Med 2011;39(05):952–960
- 16 Ide NC, Loane RF, Demner-Fushman D. Essie: a concept-based search engine for structured biomedical text. J Am Med Inform Assoc 2007;14(03):253–263
- 17 HCUP CCS. Clinical classification software ICD-9-CM. Available from: www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. Accessed February 11, 2019
- 18 Edinger T, Demner-Fushman D, Cohen AM, Bedrick S, Hersh W. Evaluation of clinical text segmentation to facilitate cohort retrieval. AMIA Annu Symp Proc 2018;2017:660–669
- 19 Thanaki J. Python Natural Language Processing. Mumbai, India: Packt; 2017
- 20 BRAT: a web-based tool for NLP-assisted text annotation. Available from: <http://aclweb.org/anthology/E/E12/E12-2021.pdf>. Accessed February 11, 2019
- 21 Unified medical language system (UMLS). Available from: <http://www.nlm.nih.gov/research/umls/>. Accessed February 11, 2019
- 22 Entrez programming utilities help. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK25501/>. Accessed February 11, 2019
- 23 Bethesda P.(MD): National Library of Medicine (US). 1946 Available from: <http://pubmed.gov/>. Accessed February 11, 2019