Bryant University

HONORS THESIS

The Impact of Doctor-Patient Communication, Health Literacy, and Mental Health Among Women with Gestational Diabetes Mellitus On Healthcare Usage

BY Marissa Stern

ADVISOR • Dr. Julie Volkman

EDITORIAL REVIEWER • Dr. Alicia Lamere

_ Submitted in partial fulfillment of the requirements for graduation
with honors in the Bryant University Honors Program
April 2023

Table of Contents

ABSTRACT

Pregnancy is a critical period in a woman's life. Overall, little research has been done on pregnant women. Pregnant women are considered a special population with special needs. All pregnant women must receive proper care from their providers and be a part of the decision-making process. Additionally, health care providers must provide materials to patients in hopes to increase their health literacy. Increasing health care usage and health literacy is meant to increase patient-centered care. This study plans to use the *All of Us* data set. We are interested in the data from the survey component of the data set. Participants respond to questions on a variety of topics, such as lifestyle, demographics, and health care. The goal of this study is to determine how much of an impact doctor-patient communication, health literacy, and mental health have on women with gestational diabetes.

INTRODUCTION

Health-related decisions are most important during pregnancy since these decisions can affect the mother and the unborn child. Pregnant mothers need access to health-related information to remain healthy and have a safe delivery. Access to this information will allow women to recognize and manage complications that arise from pregnancy, such as gestational diabetes (Kohan et al., 2007). Researchers have found that the Internet has become a popular source for pregnant women when seeking medical information (Gao, 2019). This is due to its accessibility and low cost. A Swedish study found that 84% of women relied on the Internet to provide them with health-related information while pregnant (Larsson, 2007). However, most of the information on the Internet is often outdated or incorrect (Romano, 2007). Relying on the Internet will eventually create a divide between the physician and the patient since mothers are apprehensive to share what they have read online and physicians are unaware of the incorrect beliefs mothers may have (Bert et al., 2013). According to Ghiasi (2021), pregnant women seek information regarding the fetus, nutrition, and labor and delivery. In the United States, women typically seek medical attention from various physicians. Although maternity care by family physicians has decreased in recent years, 34.4% still received care from a family physician (Kozhimannil & Fontaine, 2013). Women also receive care from obstetrician-gynecologists, and midwives. Although women primarily consulted with their health care providers, women also turned to family and friends. The same study found three issues to be the most prevalent in blocking individuals from accessing health-related information. They found that patients felt embarrassed talking about pregnancy-related issues, did not have access to accurate information, and found themselves waiting in long lines to meet with health care professionals (Ghiasi, 2021).

Gestational diabetes mellitus (GDM) is a glucose intolerance that first appears during pregnancy (Buchanan et al., 2007). Like other forms of hyperglycemia, GDM occurs when the pancreatic β-cell function does not meet the body's required amount of insulin. Clinicians detect GDM either through clinical risk assessment, glucose tolerance screening, and formal glucose tolerance screening (Buchanan et al., 2007). Physicians test for gestational diabetes between twenty-four and twenty-eight weeks of pregnancy. However, women with preexisting diabetes are not subjected to these types of tests. This type of insulin resistance may result from increased weight gain or hormones generated by the placenta. Gestational diabetes complicates 7% of all

pregnancies (American Diabetes Association, 2004). Women who suffer from GDM during pregnancy are at an increased risk of developing type 2 diabetes after childbirth (American Diabetes Association, 2004). Children of mothers whose pregnancies were complicated by GDM are at an increased risk of diabetes and obesity during young adulthood. This study aims to compare healthcare usage among women with GDM to women without GDM and measure how mental health and health literacy play a role in an individual's ability to monitor their GDM.

LITERATURE REVIEW

Doctor Patient Communication

The quality of the relationship between a doctor and his or her patient depends on the physician's communication and interpersonal skills. The better a physician's communication skills are, the more likely he or she will be able to provide an accurate diagnosis and establish positive relationships with patients (Duffy et al., 2004; Brédart et al., 2005). Establishing a caring relationship with patients will help achieve a better outcome and increase patient satisfaction (Brinkman et al., 2007). Existing studies have shown that basic communication skills alone are not enough to create a successful doctor-patient relationship (Duffy et al., 2007). The introduction of interpersonal skills helps combine both patient- and doctor-centered approaches (Brédart et al., 2005). Although, prior studies have found that physicians tend to overestimate their ability to communicate with patients. One study found that 75% of orthopedic surgeons believed that they communicated efficiently with patients, but only 21% of patients agreed that they were happy with their doctor's communication skills (Tongue et al., 2005). Patients determine whether a physician has effective communication skills by evaluating their "bedside manner" (Hall et al., 1981, p. 18).

Benefits of Effective Communication

Effective doctor-patient communication is comprised of three goals: establishing a positive relationship between doctors and patients, exchanging information, and including patients in the decision-making process (Brédart et al., 2005; Arora, 2003; Lee et al., 2002; Platt & Keating, 2007). Patients who have effective communication with their doctor are more likely to share essential information for accurate diagnosis, listen to advice, and follow through with medical treatment (Tongue et al., 2005; Arora, 2003). Good doctor-patient communication has also shown an increase in well-being and mental health (Tongue et al., 2005). Additionally, studies have shown that patients with better doctor-patient communication spend less time in the hospital, which decreases medical costs for patients (Hall et al., 1981). A more patient-centered visit also increases doctor satisfaction. Happy patients are less likely to complain or launch malpractice complaints (Hall et al., 1981). Lastly, pleased patients increase physicians' job satisfaction and reduce burnout (Brédart et al., 2005; Maguire & Pitceathly, 2002). Both job satisfaction and reduced burnout will allow for better health care delivery. The current study

aims to evaluate how doctor-patient communication determines whether women take advice and follow through with treatment for GDM.

Shared Decision Making

Scholars have defined shared decision-making (SDM) as an approach where physicians and patients share information and patients are encouraged to consider assorted options to achieve their ideal result (Elwyn et al., 2010). However, physicians are still unclear on how to implement shared decision making in the office. Elwyn et al. 2010 found that healthcare professionals doubt the benefit or practicality of implementing shared decision making. Some physicians believe that patients cannot understand medical terminology that would help them make the best decision. However, a patient's inability to understand something lies within the fact that patients are uninformed and will be unable to decide what is important to them. This supports the idea that all patients need to be better informed about key issues. In addition, others believe that patients do not want to be involved in making medical decisions and would rather have the professionals decide. Lastly, other physicians believe that implementing shared decision making into practice is impractical. Time constraints do not allow physicians to sit and explore assorted options with their patients. (Elwyn et al., 2012). Other healthcare professionals feel that they are already implementing shared decision making, but data has shown otherwise (Zikmund-Fisher et al., 2010). To be successful, shared decision making must be built on good doctor-patient communication (Stewart et al., 1995).

Shared Decision-Making During Pregnancy

Shared decision making is vital for women during pregnancy, labor, and birth. Although there is compelling evidence supporting the need for shared decision making, little has been done to implement SDM in practice. Decision aids (DA) are client-centered tools that patients can use to provide information about options and help consumers compare available options and their values (Stacey et al., 2017). Decision aids vary in form, such as paper, video, audio, or a type of interactive media, such as mobile applications. While women also need to consider what is best for themselves, they also need to decide what is best for their child.

Shared decision making is vital when it comes to the timing of delivery for women with GDM. Medical experts suggest that women with uncomplicated GDM deliver at 38 weeks (American College of Obstetricians and Gynecologists, 2013). However, past recommendations suggested earlier induction of labor. The American College of Obstetricians and Gynecologists and the Society for Maternal-Fetal Medicine now suggest that women with well-controlled gestational diabetes should not be induced before 39 weeks. It is suggested that women with poorly controlled GDM be induced before 38 + 6 days gestation. An early term or term delivery is recommended if the vascular complication is present in women with pregestational diabetes (American College of Obstetricians and Gynecologists, 2013). However, it may be difficult to implement these guidelines into practice. Obstetric, biomedical, and psychosocial factors influence the timing of delivery (Kalra et al., 2016). Therefore, it is optimal for shared decision making to occur between a woman and her physician in order to decide the best course of action.

Health Literacy

Put simply, literacy is the ability to read and write. The American 1991 National Literacy Act provides a complete and comprehensive definition, which states that individuals can read and write in English and solve problems that allow someone to function in the workplace and survive in society (National Institute for Literacy, 2008). Although illiteracy (the inability to read and write) is common in many developing countries, it is also common in developed countries as well, such as the United States. In the U.S. alone, up to fifteen million people are illiterate. This means that these individuals lack basic reading skills. In addition, as many as twenty-seven million people living in the United States lack sufficient literacy skills to fully participate in society's socioeconomic activities (Weiss et al., 1992). Individuals with low literacy often find themselves in lower-quality jobs and normally receive a lower income than someone with higher literacy. In addition, research has shown that individuals with a low literacy rate have a higher unemployment rate than those with advanced schooling. Lastly, an individual's mental health may be affected by their inability to understand essential information, therefore lowering self-esteem levels ("The Challenge: Causes of low literacy," n.d.).

Researchers have shown an increasing interest in understanding the importance of health literacy. The term 'health literacy' was first introduced in 1974. The term first came about during

a discussion of health education as a policy issue. Over time, the definition has continued to evolve. Health literacy has been defined as the degree to which individuals can obtain, understand, and communicate health-related information to make knowledgeable health decisions (Davis & McCormack, 2010). Existing research related to health literacy has examined mental health literacy and postpartum depression (Guy et al., 2014), relations between breastfeeding and maternal health literacy (Khorasani et al., 2016), and beliefs about oral health among pregnant women (Boggess et al., 2011).

Despite its introduction in the World Health Organization's Health Promotion Glossary, health literacy is still a confusing concept (Peerson & Saunders, 2009). Although health literacy relates to health promotion and preventive health, researchers find health literacy to be a difficult concept to measure and influence (Peerson & Saunders, 2009). However, many researchers believe that individuals with low health literacy do not receive preventative care, suffer from more chronic illnesses, have poorer physical and mental health function, and have higher hospitalizations (Hibbard et al., 2007). Furthermore, Green et al. (2007) show that various governments and organizations are beginning to view health literacy as an equity issue and an individual's right to citizenship. Due to this, illiteracy has been associated with poor health outcomes (Weiss et al., 1992).

Pregnancy and Health Literacy

Understanding and applying medical information is crucial throughout one's life. However, health behaviors become even more important during pregnancy, as both the mother and child can be equally affected. Even though women are confronted with various sources regarding health behavior, women with low health literacy are less likely to take folic acid during pregnancy, engage in prenatal care, and have increased hospitalizations (Song et al., 2012). In addition, they are less likely to continue breastfeeding two months postpartum (Kaufman et al., 2001). Women with low health literacy levels are less likely to understand written instructions on prenatal services and are therefore less likely to make educated health decisions (Kilfoyle et al., 2016). Improving health literacy for pregnant women is essential as their decisions will continue to affect the child years after birth.

Gestational Diabetes and Health Literacy

GDM needs to be carefully monitored by physicians and patients. However, a woman's control over gestational diabetes may be impacted by her lifestyle or health literacy level. One study showed that women who struggled with managing their gestational diabetes suffered from low health literacy (Pirdehghan et al., 2020). However, this issue was most prevalent in women that were illiterate or had minimal education, lived in rural areas, or were older housewives (Pirdehghan et al., 2020). If health literacy is improved among women, they will have better resources to adopt a healthier lifestyle. However, one major concern with women managing their GDM is that they must come to terms with their diagnosis quickly. Existing literature supports that women with GDM are crisis-oriented and may have difficulty learning how to take care of themselves (Mohamed & Ahmed, 2019). Overall, increasing a women's knowledge about GDM has the potential to improve maternal and fetal outcomes (Mohamed & Ahmed, 2019).

Mental Health

This study also seeks to analyze see how mental health varies in women with and without gestational diabetes. In general, the lifetime prevalence of depression in women is almost twice that in men (Fried et al., 2015). This may be due to genetics, hormonal fluctuations caused by the reproductive system, and psychosocial events such as gender roles. Studies also show that women are more likely to develop depression during child-bearing years (Fried et al., 2015). Additionally, other reproductive events such as infertility, miscarriage, and hormonal contraceptives may also bring on depressive episodes. The Beck Depression Inventory (BDI) states that individuals with depression feel irritable and pessimistic (Beck et al., 1996). Other depressive symptoms include insomnia, anxiety, suicidal intentions, energy loss, loss of sexual interest, and feelings of worthlessness (Fried et al., 2015). While pregnant, 18.8% of women presented anxiety symptoms before childbirth, while 20.2% of women presented anxiety symptoms after childbirth (George et al., 2013). This shows how anxiety is just as prevalent as depression during and after pregnancy. Research has shown that anxiety symptoms are prevalent in pregnant women and poor coping skills allow for anxiety symptoms to persist after childbirth. Therefore, current research must focus on finding ways to prevent mental illness, especially during the perinatal period. This study seeks to analyze how anxiety and depression levels vary in women with and without gestational diabetes.

Pregnancy and Mental Health

Little research has analyzed mental health during pregnancy, especially if the mother develops gestational diabetes. According to a study put forth by Ko et al. (2012), only 40% of women are treated for postpartum depression. Providing women with the proper skills to recognize and manage symptoms may decrease the effects that are felt on women, babies, and their family members. One study found that two conditions increased a women's chance of developing postpartum depression: high levels of stress and being in a lower socioeconomic bracket (McIntosh, 1993). Existing literature reveals that women were often unable to recognize that they were suffering from postpartum depression and were only made aware of it once another person informed them that they were not behaving like their usual selves (Abrams et al., 2009; Letourneau et al., 2007). Women often claim that their symptoms arise from caring for a newborn alone, the inability to breastfeed or have a vaginal delivery, lack of sleep, postpartum period, or having no place to live or work (McIntosh, 1993; Ugarizza, 2002). Even though some women were able to recognize depressive symptoms in themselves, many were apprehensive to seek medical treatment. This was because women worried that medical professionals would share this information with their child's school and be seen as unfit mothers, giving authorities the right to take custody of their children (Guy et al., 2014; McIntosh, 1993).

Research also supports that depression in pregnant women more often went undiagnosed than in non-pregnant women, even though pregnant women more routinely interact with the health care system (Ko et al., 2012). This supports that medical professionals frequently miss opportunities to identify mental illness in patients and connect them with the proper resources. In a study of 9,028 women, researchers found that depressive symptoms were more common during pregnancy than after pregnancy (Evans et al., 2001). Women experiencing depressive symptoms while pregnant were at a higher risk of developing postpartum depression and having adverse birth outcomes. Early detection (i.e., during pregnancy) has been shown to decrease harmful effects on the mother and baby during pregnancy and reduce postpartum mood and anxiety disorders. Although antenatal depression is treatable, most women are reluctant to take antidepressants during pregnancy (Jeong et al., 2013). Since medications can have negative effects on the fetus, research studies whether there is a correlation between emotional support and antenatal depression (Jeong et al., 2013). Results show that pregnant women lean on their

partner and mother for emotional support. In addition, past research has also shown that current emotional support is a better determinant of developing antenatal depression than past support (Jeong et al., 2013).

Mental Health and Gestational Diabetes

Existing literature supports the notion that a woman's quality of life decreases in both short and long terms when diagnosed with gestational diabetes (Marchetti et al., 2017). However, this does not necessarily mean that the condition directly affects a woman's quality of life. An underlying psychological element is deemed to be the culprit. In most cases, a woman's quality of life decreases when she is not equipped with the proper resources to respond to GDM symptoms (Marchetti et al., 2017). Research has shown that women had increased levels of anxiety when initially diagnosed with GDM but were consistent with those of women who were glucose tolerant only weeks later (Daniells et al., 2003). By week thirty-six, there were no differences in anxiety levels between both groups of women. However, there is little research that determines whether an individual's mental health plays a role in how she manages her gestational diabetes. This study will bridge the gap between mental health and healthcare usage among women with gestational diabetes.

CURRENT STUDY

Rationale

Past research has primarily focused on the importance of doctor-patient communication, health literacy, and mental health (Duffy et al., 2004; Song et al., 2012; Fried et al., 2015). However, little research has been done to analyze how these mediators impact healthcare usage among women with GDM to women without GDM. This study is seeking to analyze how doctor-patient communication, health literacy, and mental health impact how women with and without GDM interact with the healthcare system.

Research Questions

RQ1: How do pregnant women with GDM differ from those without GDM on mental health (anxiety and depression)?

RQ2: How do pregnant women with GDM differ from those without GDM on their healthcare usage?

RQ3: How do pregnant women with GDM differ from those without GDM on health literacy?

RQ4: How do pregnant women with GDM differ from those without GDM on shared decision making communication with their provider?

METHODS

Data Set

This study will use the *All of Us* Research Program Data Center (DRC) led by Vanderbilt University Medical Center, sponsored by the National Institutes of Health (NIH) (National Health Institute, n.d.). As of March 24, 2022, approximately 477,000 individuals have participated in *All of Us. All of Us* does not report the exact number of individuals participating in the study to protect patients' privacy and confidentiality. Currently, 59.5% of participants identify as female. The age of participants ranged from 18- 89+ years. The data shows that 21.6% of participants are between the ages of 60-69 years, 19.3% are between 50-59 years, 14.8% are between the ages of 70-79, and only 10.2% of participants are between the ages of 18-29 years. The ethnic/race profile of the data set is 48.5% White, 20.4% Black, African American, or African, 16.9% Hispanic or Spanish, and 6.6% more than one race/ethnicity.

The total number of pregnant women registered with the All of Us Dataset is 6,969. There was 1,031 women with GDM and 5,939 women without GDM. Regarding race and ethnicity, we found that a majority of the sample was Caucasian. Regarding women without GDM, over 2,000 individuals were Caucasian. A sample of just under 2,000 individuals did not indicate their race and there was just over 1,000 women who specified that they were Black or African American. We also that a very small sample of the population were Asian or from another single population. Regarding women with GDM, a majority of the respondents did not indicate their race. We found that a little under 300 respondents were Caucasian and under 300 were Asian or Black or African America.

Overall, using the All of Us dataset presented some challenges. We found that the database was difficult to utilize for several reasons. For one, it was not clear on how to use the database. It was not an intuitive dataset and Meaning, it was difficult to use the variables to answer the research questions. It was difficult to make cohorts on the All of Us site and it was not clear on how to assign different variables to different groups. Everyone had access to the project, but it was difficult for everyone to view it at all times. We also found that running the code in the cloud that All of Us created was not as easy as we had hoped. All of Us did host weekly office hours, but it seemed that there was no direct person to contact with any questions.

Sampling Method

The *All of Us* Research Program collects data from a wide variety of sources. These include surveys, electronic health records (EHRs), biosamples, physical measurements, and wearables, like Fitbit. Participants are invited to enroll by partner health care provider organizations (i.e., academic medical centers), Veteran Affair medical centers, and community health centers. Participants also have the option to enroll on their own by visiting JoinAllofUs.org or by attending *All of Us* events. Thus, the *All of Us* Research Program follows a convenient, non-profitability sampling method where participants are selected based on their convenient accessibility.

Participants have the option to share various kinds of information by completing surveys, granting access to their electronic health records (EHRs), and syncing Fitbit devices with the *All of Us* portal. Some participants are invited to visit the partner site to have physical measurements and blood and urine samples taken. *All of Us* continues to stay in contact with participants about new opportunities to share data through additional surveys, new research studies, and new electronic tools, including apps.

HER Component of Data Set

All participants were invited to share their electronic health records (EHRs) with the *All of Us* Research Program. *All of Us* removes identifiers from participants' EHR data before adding this information to be available for research. *All of Us* used the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) to standardize all EHR data. For this study, the EHR data related to pregnancy, gestational diabetes mellitus, and anxiety and depressive disorders.

Survey Component of Data Set

This study is interested in the participants that answered the survey portion of the data set. Other components of the data set include information from the electronic medical record and biosamples. Participants in the *All of Us* Research Program respond to surveys that cover a variety of topics, including demographics, health care, and lifestyle. Each survey has been assessed for readability and accessibility. *All of Us* use cognitive interviews and quantitative testing. To capture a sample that reflects the U.S population, this testing process includes people

from different educational backgrounds and geographic locations. Participants can complete additional surveys on health care access, personal and family medical history, and other topics. The purpose of this study is interested in the Health Care Access and Utilization and Overall Health sections of the survey.

MEASURES

Independent Variable

This study is specifically interested in pregnant women either with or without gestational diabetes mellitus. From the electronic health records (EHR), we found that 1,480 women were diagnosed with gestational diabetes mellitus. The *All of Us* dataset found that 2% or 900 women were between the ages of 30 and 39 years of age. The dataset also shows that 460 participants were between the ages of 18 and 29 and 220 participants were between the ages of 40 and 49. No participants were 50 years or older. The count breakdown was divided into GDM in childbirth, postpartum GDM, GDM complicating pregnancy, GDM class A>1<, and GDM class A>2<. While GDM A1 can be controlled with diet and exercise, patients need medication to help manage their GDM A2.

Dependent Variables

Healthcare Usage

The dependent variable of the present study is health care usage and shared decision-making. To determine health care usage, this study looked at the Health Care Access and Utilization Survey. Each question was generated by a group of panelists from *All of Us.* The first question asked whether a participant had seen or talked with a doctor who specializes in women's health in the last twelve months. This question required a yes or no answer.

Shared Decision Making

This study also looked at the Health Care Access and Utilization Survey when looking at shared decision-making. The first question asked whether participants felt that they were treated with respect by their doctors or health care providers. The response scale ranged from "None of the Time" to "Always." The second question asked whether health care providers asked for participants' opinions or beliefs about their medical care or treatment. The response scale ranged from "None of the Time" to "Most of the Time." The third question asked how often participants' doctors or health care providers shared information about their health and health care that was easy to understand. The response scale ranged from "None of the Time" to "Always."

Moderators

Mental Health

To understand the impact of mental health on health behaviors, this study also accounted for the impact of depression and anxiety disorders. This information was generated from the EHR. This information was found under the Conditions domain of the data browser. For anxiety, the count breakdown was divided into anxiety disorder, anxiety state, panic, and anxiety about body function or health. Of the participants registered with *All of Us,* 39,220 women suffered from anxiety. Regarding depression, the count breakdown was divided into major depressive disorder, recurrent depression, moderate depression, and postpartum depression. The *All of Us* dataset found that 40,700 women suffered from depression.

Health Literacy

This study quantified patient health literacy using three questions and their responses from the Overall Health Survey. The first question asked how confident patients were in filling out medical forms on their own. The response scale ranged from "Not at All" to "Extremely." The second question asked how often someone receives help reading health-related materials. The response scale ranged from "Always" to "Never." The final question asked how often someone has problems learning about their medical condition because of difficulty understanding written materials. The response scale ranged from "Always" to "Never." Each question was generated by panelists from *All Of Us.*

DATA ANALYSIS

Data analysis will be conducted using R. Preliminary analysis will include Chi Square tests and Mann-Whitney U tests. The Chi Square test will be used to test anxiety and depression levels among women with and without GDM. The Mann-Whitney U test will be used to see if there is an association between each category and the data responses. This study will follow all procedures for large sample sizes. is to see whether there is a significant relationship between moderators and dependent variables.

ETHICAL CONSIDERATIONS

All of Us is a publicly available data set. The research program received IRB approval from Vanderbilt University Medical Center. All privacy protocols will be followed per the *All of Us* Data Use Agreement policies. The *All of Us* Research Program requires that researchers' complete ethics training in which they must agree to several rules. One rule is that researchers will not try to find out who the participants are. Privacy is extremely important to the research program. The database removes names and other identifying information, and they have Certificates of Confidentiality from the U.S. government.

RESULTS

Throughout this study, we have looked at various factors. These factors include healthcare usage, health literacy, mental health (anxiety and depression), and shared decision making with a provider. The first research question analyzed how pregnant women with GDM differ from those without GDM on mental health. The second research question analyzed how women with GDM differ from those without GDM on healthcare usage. The third research question analyzed how women with GDM differ from those without GDM on health literacy. The fourth research question analyzed how pregnant women with GDM differ from those without GDM on shared decision making communication with a healthcare provider.

When analyzing the data, we decided to merge and not merge the survey responses. When merging the survey responses, we combined similar responses (i.e., sometimes/often) to test whether that affected potential significance. Unmerged responses are when we treated each response separately.

RQ1 Testing Independence of Anxiety/Depression and GDM Status

Regarding anxiety, a Pearson's Chi-square test with Yates' continuity correction was performed. The null hypothesis (that anxiety and GDM are independent) was rejected. The results show that fewer people than would be expected have anxiety and GDM. Meaning, women without GDM were more likely to experience symptoms of anxiety, X-squared = 464.65, df = 1, p-value < 2.2e-16. See Appendix A.

Regarding depression, a Pearson's Chi-squared test with Yate's continuity correction was performed. Once again, we reject the null hypothesis that depression and GDM are independent. The results support that fewer people than would be expected have depression and GDM, X-squared = 41.053, df = 1, p-value = 1.482e-10. See Appendix B.

RQs 2-4 Testing Association between Utilization Survey and Overall Health Survey Results and GDM Status

The Healthcare and Utilization Survey portion of All of Us included questions related to Research Questions 2-4 which asked whether pregnant women with GDM differ from pregnant women without GDM on their health usage, health literacy and shared decision making. We decided to merge and not merge question responses to see if there were differences in reporting

significance. The Overall Health survey portion of All of Us included questions related to Research Question 3, which asked about health literacy.

RQ2

The second research question focused on healthcare usage. Coming from the Utilization Survey, the first question asked if an individual has visited or talked to a doctor who specializes in women's health. A Mann Whitney U Test was performed. The results showed that there was no association between an individual's response and their GDM status (either having GDM or not having GDM).

RQ3

The third research question focused on health literacy. Survey questions came from Healthcare and Utilization Survey. The first question asked how often a doctor shared health care information that was easy to understand. The merged data rejected the null hypothesis, W = 513604, p-value = 0.01761. The unmerged data failed to reject the null hypothesis, W = 5244436, p-value = 0.3674.

The Overall Health Survey in All of Us included questions related to Research Question 3 as well.

The first survey question asked whether individuals were confident in filling out medical forms by themselves. When merging the question responses, we found that there was a possible rejection of the null hypothesis. The alternative hypothesis is that the true location shift is not equal to 0. The results showed a non-significant relationship between women with GDM and women without GDM in filling out medical forms by themselves, W = 6694588, p-value = 0.04667. In not merging the question responses, we failed to reject the null hypothesis, W = 6861617, p-value = 0.6433.

The second survey question asked how often an individual had help reading health-related materials. The merged data failed to reject the null hypothesis, W = 6410495, p-value = 0.133. The unmerged data also failed to reject the null hypothesis, W = 6662518, p-value = 0.1528. The third survey question asked how often an individual had problems learning about a medical

condition due to the difficulty of understanding written information. The merged data showed that it was possible to reject the null hypothesis, W = 6329878, p-value= 0.1013. The unmerged data showed that it was possible to reject the null hypothesis, W = 6639532, p-value = 0.07801.

RQ4

The fourth research question focused on shared decision making communication between the patient and the provider. The first survey question asked if an individual was treated with respect by her health care providers. Two Mann Whitney U Tests were performed. In merging the question responses, we failed to reject the null hypothesis, W = 504266, p-value = 0.879. The true location shift is not equal to 0. When the data was not merged, we also failed to reject the null hypothesis, W = 532764, p-value = 0.2869. The second survey question asked whether medical professionals asked for an individual's opinions or beliefs regarding medical care or treatment. The merged data failed to reject the null hypothesis, W = 518677, p-value = 0.9924. The true location shift is not equal to 0. The unmerged response questions also failed to reject the null hypothesis, W = 547777, p-value = 0.4357.

DISCUSSION

The purpose of this study was to test whether women with GDM differ in shared-decision making, health literacy, and mental health compared to women without GDM. Health related decisions are incredibly important, especially during pregnancy. In order to have a safe delivery, pregnant women need to have access to up to date and reliable information. However, further research needs to be completed to see whether there are differences in pregnant women with GDM and those without. If so, materials and techniques need to be adapted to better help each population. The first research question asked whether pregnant with GDM differ from those without GDM on mental health (anxiety and depression). The second research question asked whether pregnant women with GDM differ those without GDM on their healthcare usage. The third research question looked to see whether women with GDM differ from those without GDM on health literacy. The fourth research questions looked at whether women with GDM differ from those without GDM on shared decision making with their provider.

Little research has been done on pregnant women, especially those diagnosed with GDM. It is extremely important for pregnant women to be a part of the decision making process and receive the proper care necessary. The differences in women with GDM to those without remain unknown. Identifying health literacy levels, mental health struggles, and patient provider communication will facilitate the development of researching tools that will allow women to feel more comfortable with taking control of their health care.

Through utilizing the All of Us database, we analyzed several questions related to health literacy, mental health (anxiety and depression) and doctor-patient communication. We adapted questions from the All of Us database to answer our research questions of interest. While we found several non-significant results in our analyses, there were some suggestions of significant relationships to explore.

For the first research question, we tested whether women with GDM faced higher levels of anxiety and depression than women without GDM. The results were significant, though not in the way that we had expected. The results showed that women without GDM had higher levels of anxiety and depression than women with GDM. This may be due to the fact that the All of Us

data set has a small percentage of women with GDM. These findings may not be representative of the entire GDM population.

The second research question focused on healthcare usage and whether an individual has visited or spoken to a doctor who specializes in women's health. The results were non-significant. From this, we can interpret that all are talking to a doctor who specializes in women's health. Additionally, the findings suggest that visiting a doctor who specializes in women's health is relatively easy to do and gynecologists are available to any individual who would like to meet with them.

The third research question focused on health literacy. The question from the Healthcare and Utilization Survey asked whether doctor's shared information that easy to understand with their patients. When merging the data responses, it is suggested that pregnant women with GDM and those without GDM are different in receiving health information from their doctor that is easy to understand. However, when unmerging the data, this significance is not seen. For the Overall Health questions in the All of Us dataset, the first question asked how confident individuals felt in filling out medical forms by themselves. However, the results were nonsignificant. This suggests that all women, regardless of GDM status, felt confident in filling out medical forms by themselves. Though, the merged data responses does subject a difference could be valued, if we adjusted the cut-off of significance to $p = .10$. The second question asked how often individuals had help reading health-related materials. The results were non-significant and it is suggested that pregnant women with or without GDM have no difference in problems reading health materials. The responses suggest that both are doing well in reading health-related materials. The last question. This may be because the sampled came from a higher socioeconomic status and these individuals typically have a higher level of health literacy.

The fourth research question focused on shared decision making between the patient and the provider. The first survey question asked if an individual felt that she was treated with respect by a health care provider. Unfortunately, there were no significant differences in responses when merging and not merging the data. This suggests that all women, regardless of GDM status, were treated with respect by their provider. The last question asked if individuals were asked for their opinions or beliefs regarding medical care. Neither merging the data or not merging the data

provided significant results. The data suggests that pregnant women are being asked their opinions and beliefs regarding care, no matter their GDM status.

Women with GDM may differ from women without GDM for several reasons. Prior research has found that women who struggled with managing their GDM also suffered from low health literacy (Pirdehghan et al., 2020). Additionally, researchers have found that women with GDM are crisis-oriented and may have difficulty taking care of their gestational diabetes. Meaning, proper resources need to be produced to ensure that women are able to effectively manage their GDM.

However, the data infers that doctor-patient relationships are continuing to improve and movements toward patient-centered care is working. Additionally, it can be inferred that health literacy efforts are being implemented regardless of an individual's GDM status. This suggests that people are receiving information in a way that they can understand. Even though we are currently seeing a positive outcome, continued research is necessary to ensure that materials continue to improve.

LIMITATIONS

This study faced several limitations. First, we were unable to generate our own questions and had to use what the database provided. Since the database had predetermined questions, we were unable to ask more specific questions that were geared towards our research questions. Second, there were no open-ended questions. Due to this, we were unable to get more in-depth responses. Generally, the question options ranged from "none of the time" to "always." Third, there was a lack of reliable data. It is unclear whether all the participants were honest in their responses. Fourth, many of the participants are Caucasian. Due to this, we were unable to get a comprehensive read on the entire population of women with GDM. There are several reasons why this is harmful. Research has found that women of color have complained of more difficulty in seeing a specialist when needed (Wyn et al., 2004). This is typically since health insurance coverage varies considerably among different populations of women. Additionally, it was noted that women of color tend to have more diffculty finding the time to visit the doctor. Research has also shown that African American women and Latinas are more likely to receive care at clinics as opposed to in a doctor's office (Wyn et al., 2004). Databases should be more aware of how inclusive they are being. The All of Us Database tends to recruit individuals through the patient's doctor's office and if women of color are having difficulty attending a doctor, there is not way for them to join the program.

Additionally, more women did not have GDM. This influenced how many women with GDM we were able to access. Lastly, there was no longitudinal data. We saw that during pregnancy women had high literacy levels, but we do not know how this compared to before they were pregnant. Meaning, we were unable to tell if they had high literacy levels before becoming pregnant or acquired those skills during pregnancy.

CONCLUSION

Overall, this study shows that more research is needed. More research needs to be conducted on women with GDM. Through this study, we learned that most women were able to take better care of themselves during pregnancy. However, this study did not look at many women of color. Future research may include interviewing women with GDM from various backgrounds or sending out surveys. Additional research may want to examine different questions related to health literacy and doctor-patient communication to better understand what is happening during medical encounters.

REFERENCES

Abrams, L. S., Dornig, K., and Curran, L. (2009). Barriers to service use for postpartum depression symptoms among low-income ethnic minority mothers in the United States. *Qualitative Health Research, 19*(4), 535-551. Doi:10.117.1049732309332794

All of US public data browser. All of Us Public Data Browser. (n.d.). Retrieved April 5, 2022, from https://databrowser.researchallofus.org/

American College of Obstetricians and Gynecologists. (2013). "ACOG committee opinion no. 506: Medically indicated late-preterm and early-term deliveries." *Obstet Gynecol, 121*(4), 908-910

American Diabetes Association. (2004). Gestational Diabetes Mellitus. *Diabetes Care, 27,* 88-90. doi:10.2337/diacare.27.2007.s88

Arora, N. (2003). Interacting with cancer patients: the significance of physicians' communication behavior. *Soc Sci Med. 57*(5), 791-806

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory (BDI-II)* (Vol. 10, p. s15327752jpa6703_13). London, UK: Pearson.

Bert, F., Gualano, M. R., Brusaferro, S., De Vito, E., de Waure C., La Torre, G., et al. (2013). Pregnancy e-health: A multicellular Italian cross-sectional study on internet use and decision-making among pregnant women. *J Epidemiol Community Health. 67,* 1013-1018

Boggess, K. A., Urlaub, D. M., Moos, M.-K., Polinkovsky, M., El-Khorazaty, J., & Lorenz, C. (2011). Knowledge and beliefs regarding oral health among pregnant women.

The Journal of the American Dental Association, *142*(11), 1275–1282. https://doi.org/10.14219/jada.archive.2011.0113

Brédart, A., Bouleuc, C., and Dolbeault, S. (2005). Doctor-patient communication and satisfaction with care in oncology. *Curr Opin Oncol. 17*(14), 351-354

Brinkman, W. B., Geraghty, S. R., Lanphear, B. P., et al. (2007). Effect on multisource feedback on resident communication skills and professionalism: a randomized controlled trial. *Arch pediatr Adolesc. 161*(1), 44-49

Buchanan, T., Xiang, A., Kjos, S., and Watanabe, R. (2007). What is Gestational Diabetes? *Diabetes Care, 30,* 105-111. doi:20.2337/dc07-s201

Daniells, S. Grenyer, B. F. S., Davis, W., Coleman, K., Burgess, J., and Moses, R. (2003). Is a diagnosis associated with an increase in maternal anxiety and stress in the short and intermediate term? *Diabetes Care, 26*(2), 385-389. doi:10.2337/diacare.26.2.385

Duffy, F. D., Gordon, G. H., Whelan, G., et al. (2004). Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med. 79*(6), 495-507

Elwyn, G., Coulter, A., Laitner, S., Walker, E., Watson, P., and Thomson, R. (2010). Implementing shared decision making in the NHS. *Bmj, 341*

Elwyn, G., Frosh, D., Thomson, R., Joseph-Williams, N., Llyod, A., et al. (2012). Shared Decision making: a Model for Clinical Practice. *Journal of General Internal Medicine. 27*, 1361-1367

Fried, E. I., Espskamp, S., Nesse, R. M., Tuerlinckx, F., and Borsboom, D. (2016). What are

'good' depression symptoms? Comparing the centrality of DSM and non-DSM symptoms of depression in a network analysis. *Journal of Affective Disorders, 189,* 314-320

Gao, L. L., Larsson, M., and Luo. S. Y. (2013). Internet use by Chinese women seeking pregnancy-related information. *Midwifery. 29,* 730-735

George, A., Luz, R. F., Tychey, C. D., Thilly, N., and Spitz, E. (2013). Anxiety symptoms and coping strategies in the perinatal period. *BMC Pregnancy and Childbirth, 13,* 233

Ghiasi, A. (2021). Health information needs, sources of information, and barriers to accessing health information among pregnant women: a systemic review of research. *The Journal of Maternal-Fetal and Neonatal Medicine. 34*(8), 1320-1330 Doi: 10.1080/14767058.2019.1634685

Green, J., (2007). Health Literacy: Terminology and trends in making and communicating health-related information. *Health Issues, 92,* 11-14

Guy, S., Sterling, B. S., Walker, L. O., & Harrison, T. C. (2014). Mental Health Literacy and Postpartum Depression: A Qualitative Description of Views of Lower Income Women. *Archives of Psychiatric Nursing*, *28*(4), 256–262. https://doi.org/10.1016/j.apnu.2014.04.001

Hall, J. A., Roster, D. L., R & C. S. (1981). Communication of affect between patient and physician. *J Health Soc Behav. 22*(1), 18-30

Hibbard, J. H., Peters, W. E., Dixon, A. and Tusler, M, (2007). Consumer competencies and the use of comparative quality information: It isn't just about literacy. *Patient Care Research*

and Review, 64, 379-394

Jeong, H., Lim, J., Lee, M., Kim, S., Jung, I., and Joe, S. (2013). The association of psychosocial factors and obstetric history with depression in pregnant women: focus on the role of emotional support. *General Hospital Psychiatry, 35*(4), 354-358

Kalra, B., Gupta Y., and Kalra, S. (2016). Timing of Delivery in Gestational Diabetes Mellitus: Need for Person-Centered, Shared Decision-Making. *Diabetes Therapy, 7,* 169-174

Kaufman, H., Skipper, B., Small, L., Terry, T., and McGrew, M. (2001). Effect on literacy of breast-feeding outcomes. *South. Med. J, 94,* 293-296

Khorasani, E., Peyman, N., & Esmaily, H. (2017). Relations Between Breastfeeding Self-efficacy and Maternal Health Literacy Among Pregnant Women *Evidence Based Care*, *6*(4), 18–25. https://doi.org/20.22038/ebcj.2016.7986

Kilfoyle, K. A., Vitko, M., O'Conor, R., and Bailey, S. C. (2016). Health Literacy and Women's Reproductive Health: A Systematic Review. *J. Women's Health, 25,* 1237-1255

Ko, J. K., Farr, S. L., Dietz, P. M., and Robins, C. L. (2012). Depression and treatment among U.S. pregnant and nonpregnant women of reproductive age, 2005-2009. *Journal of Women's Health, 21*(8), 830-836. doi:10.1089.jwh.2011.3466

Kohan, S. H., Ghasemi, S., and Dodangeh, M. (2007). Associations between maternal health literacy and prenatal care and pregnancy outcome. *Iran J Nurs Midwifery Res. 12,* 146-52

Kozhimannil, K. and Fontaine, P. (2013). Care From Family Physicians Reported by Pregnant

Women in the United States. *Ann Fam Med. 11*(4), 350-354. Doi: 10.1370/afm.1510

Larrson, M. A. (2007). A descriptive study of the use of the internet by women seeking pregnancy-related information. *Midwifery. 25,* 14-20

Lee, S. J., Back, A. L., Block, S. D., and Stewart, S. K. (2002). Enhancing physician-patient communication. *Hematology Am Soc Hematol Educ Program. 1,* 464-483

Letourneau, N. L., Duffet-Leger, L., Stewart, M., Hegadoren, K., Dennis, C., Rinaldi, C. M., et al. (2007). Canadian mothers' perceived support needs during postpartum depression. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 36*(5), 441-449. doi:10.1111/j.1552-6909.2007.00174.x.

Maguire, P., and Pitceathly, C. (2002). Key communication skills and how to acquire them. *BMJ 325*(7366), 697-700

Marchetti, D., Carrozzino, D., Fraticelli, F., Fulcheri, M., and Vitacolonna, E. (2017). Quality of Life in Women with Gestational Diabetes Mellitus: A Systematic Review. *Journal of Diabetes Research, 17,* doi: 10.1155/2017.7058082

McIntosh, J. (1993). Postpartum depression: Women's help-seeking behavior and perceptions of cause. *Journal of Advanced Nursing, 18*(2), 178-184. doi:10.1046/j.1365-2648.1993.1802178.x.

Mohamed, S. H. and Ahmed, A. A. (2019). Educational program for Health Literacy among Pregnant Women with Gestational Diabetes: its effect on Maternal & fetal outcomes. *International Journal of Nursing Didactics, 9*(4). Doi: https://doi.org/10.5520/ijnd.v9io4.2554

National Institute for Literacy. (2008) National Literacy Act 1991. www.nifl.gov/publi-law.html

Pirdehghan, A., Eslahchi, M., Esna-Ashari, F., and Borzouei, S. (2020). Health literacy and diabetes control in pregnant women. *Journal of family medicine and primary care, 9*(2), 1048-1052. doi:10.4103/jfmpc.jfmpc_891_19

Platt, F. W. and Keating, K. N. (2007). Differences in physician and patient perceptions of uncomplicated UTI symptom severity: understanding the communication gap *Int J Clin Prac. 61*(2), 303-308

Romano, A. M. (2007). A changing landscape: Implications of pregnant women's Internet use for childbirth educators. *J Perinat Educ. 16,* 18-24

Stacey, D., Legare, F., Lewis, K., Barry, M. J., Bennett, C. L., Eden, K. B., Holmes-Rovner, H., Llewellyn-Thomas, A., Lyddiatt, R., and Thomson, R. (2017). Decision Aids for People Facing Health Treatment Or Screening Decisions *Cochrane database of systemic Reviews,* (4).

Song, F. W., West, J. E., Lundy, L., and Dahmen, N. S. (2012) Women, Pregnancy, and Health Information Online: The Making of Informed Patients and Ideal Mothers. *Gend. Soc, 26,* 773-798

Spong, C. Y. (2013). Defining "term" pregnancy: Recommendations from the Defining "Term" Pregnancy Workgroup. *JAMA, 309*(13), 2445–2446.

Stewart, M., Brown, J., Weston, W., McWinney, I., McWilliam, C., and Freeman, T. (1995). Patient Centred Medicine: Transforming the Clinical Method. CRC press.

The Challenge: Causes of low literacy. Literacy Pittsburgh. (n.d.). Retrieved from https://www.literacypittsburgh.org/the-challenge/

Tongue, J. R., Epps, H. R., and Forese, L. L. (2005). Communication skills for patient-centered care: research-based, easily learned techniques for medical interviews that benefit orthopedic surgeons and their patients. *J Bone Joint Surg Am. 87,* 652-658

Ugarriza, D. N. (2002). Postpartum depressed women's explanation of depression. *Journal of Nursing Scholarship, 34*(3), 227-233. doi:10.1111.j.1547-5069.2002.00227.x.

Weiss, B., Hart, G., McGee, D., & D'Estelle, S. (1992). Health Status of Illiterate Adults: Relations Between Literacy and Health Status Among Persons With Low Literacy Skills. *JABFM*, *5*(3). https://doi.org/10.3122/jabfm.5.2.257

Wyn, R., Ojeda,V., Rnaji, U., Salganicoff, A. (2004). Racial and Ethnic Disparities in Women's Health Coverage and Access To Care Findings from the 2001 Kaiser Women's Health Survey. Retrieved from: https://www.kff.org/wp-content/uploads/2013/01/racial-and-ethnic-disparities-in-women-s-health-coverage-and-access-to-care.pdf

Zikmund-Fisher, B. J., Couper, M. P., Singer, E., Levin, C. A., Fowler, F. J., Ziniel, S., et al. (2010). The DECISIONS study: a nationwide survey of United States adulting regarding 9 common medical decisions. *Med Decis Making. 20*

APPENDICES

Appendix A – Anxiety Chart

Anxiety

GDM	Y	N
Y	308	729
N	3936	2096

Appendix B – Depression Chart

Depression

GDM	Y	N
Y	55	982
N	732	5300

Appendix C – Code from All of Us
To pull datasets from All Of Us:

```
library(tidyverse)
library(bigrquery)
```

```
# This query represents dataset "DataSet_10_13" for domain "condition" and was generated for
All of Us Registered Tier Dataset v5
dataset_68612580_condition_sql <- paste("
    SELECT
        c_occurrence.person_id,
        c_occurrence.condition_concept_id,
        c_standard_concept.concept_name as standard_concept_name,
        c_standard_concept.concept_code as standard_concept_code,
        c_standard_concept.vocabulary_id as standard_vocabulary,
        c_occurrence.condition_start_datetime,
        c_occurrence.condition_end_datetime,
        c_occurrence.condition_type_concept_id,
        c_type.concept_name as condition_type_concept_name,
        c_occurrence.stop_reason,
        c_occurrence.visit_occurrence_id,
        visit.concept_name as visit_occurrence_concept_name,
        c_occurrence.condition_source_value,
        c_occurrence.condition_source_concept_id,
        c_source_concept.concept_name as source_concept_name,
        c_source_concept.concept_code as source_concept_code,
        c_source_concept.vocabulary_id as source_vocabulary,
        c_occurrence.condition_status_source_value,
        c_occurrence.condition_status_concept_id,
        c_status.concept_name as condition_status_concept_name
    FROM
        ( SELECT
            *
        FROM
            `condition_occurrence` c_occurrence
        WHERE
            (
                condition_concept_id IN  (
                    SELECT
                        DISTINCT c.concept_id
                    FROM
                        `cb_criteria` c
                    JOIN
                        (
                            select
                                cast(cr.id as string) as id
                            FROM
                                `cb_criteria` cr
                            WHERE
                                concept_id IN (
```

```
                4024659, 441542, 4152280
            )
            AND full_text LIKE '%_rank1]%'
        ) a
            ON (
                c.path LIKE CONCAT('%.',
            a.id,
            '.%')
            OR c.path LIKE CONCAT('%.',
            a.id)
            OR c.path LIKE CONCAT(a.id,
            '.%')
            OR c.path = a.id)
        WHERE
            is_standard = 1
            AND is_selectable = 1
        )
)
AND (
    c_occurrence.PERSON_ID IN (
        SELECT
            distinct person_id
        FROM
            `cb_search_person` cb_search_person
        WHERE
            cb_search_person.person_id IN (
                SELECT
                    criteria.person_id
                FROM
                    (SELECT
                        DISTINCT person_id,
                        entry_date,
                        concept_id
                    FROM
                        `cb_search_all_events`
                    WHERE
                        (
                            concept_id IN (
                                SELECT
                                    DISTINCT c.concept_id
                                FROM
                                    `cb_criteria` c
                                JOIN
                                    (
                                        select
```

```
                                    cast(cr.id as string) as id
                                FROM
                                    `cb_criteria` cr
                                WHERE
                                    concept_id IN (4299535)
                                    AND full_text LIKE '%_rank1]%'
                            ) a
                                ON (
                                    c.path LIKE CONCAT('%.',
                                a.id,
                                '.%')
                                OR c.path LIKE CONCAT('%.',
                                a.id)
                                OR c.path LIKE CONCAT(a.id,
                                '.%')
                                OR c.path = a.id)
                            WHERE
                                is_standard = 1
                                AND is_selectable = 1
                            )
                            AND is_standard = 1
                        )
                    ) criteria
                ) ))
            ) c_occurrence
        LEFT JOIN
            `concept` c_standard_concept
                ON c_occurrence.condition_concept_id = c_standard_concept.concept_id
        LEFT JOIN
            `concept` c_type
                ON c_occurrence.condition_type_concept_id = c_type.concept_id
        LEFT JOIN
            `visit_occurrence` v
                ON c_occurrence.visit_occurrence_id = v.visit_occurrence_id
        LEFT JOIN
            `concept` visit
                ON v.visit_concept_id = visit.concept_id
        LEFT JOIN
            `concept` c_source_concept
                ON c_occurrence.condition_source_concept_id =
c_source_concept.concept_id
        LEFT JOIN
            `concept` c_status
                ON c_occurrence.condition_status_concept_id = c_status.concept_id", sep="")
```

```
condition_68612580_path <- file.path(
  Sys.getenv("WORKSPACE_BUCKET"),
  "bq_exports",
  Sys.getenv("OWNER_EMAIL"),
  "condition_68612580",
  "condition_68612580_*.csv")
message(str_glue('The data will be written to {condition_68612580_path}. Use this path when
reading ',
                 'the data into your notebooks in the future.'))

bq_table_save(
  bq_dataset_query(Sys.getenv("WORKSPACE_CDR"), dataset_68612580_condition_sql,
billing = Sys.getenv("GOOGLE_PROJECT")),
  condition_68612580_path,
  destination_format = "CSV")

read_bq_export_from_workspace_bucket <- function(export_path) {
  col_types <- NULL
  bind_rows(
    map(system2('gsutil', args = c('ls', export_path), stdout = TRUE, stderr = TRUE),
        function(csv) {
          message(str_glue('Loading {csv}.'))
          chunk <- read_csv(pipe(str_glue('gsutil cat {csv}')), col_types = col_types,
show_col_types = FALSE)
          if (is.null(col_types)) {
            col_types <- spec(chunk)
          }
          chunk
        }))
}
dataset_68612580_condition_df <-
read_bq_export_from_workspace_bucket(condition_68612580_path)

library(tidyverse)
library(bigrquery)

# This query represents dataset "DataSet_10_13" for domain "survey" and was generated for All
of Us Registered Tier Dataset v5
dataset_68612580_survey_sql <- paste("
    SELECT
        answer.person_id,
        answer.survey_datetime,
        answer.survey,
        answer.question_concept_id,
        answer.question,
```

```
    answer.answer_concept_id,
    answer.answer,
    answer.survey_version_concept_id,
    answer.survey_version_name
  FROM
    `ds_survey` answer
  WHERE
    (
      question_concept_id IN (
        43530439, 1585772, 1585778, 43530401, 1585766, 43530438, 43530437
      )
    )
    AND (
      answer.PERSON_ID IN (
        SELECT
          distinct person_id
        FROM
          `cb_search_person` cb_search_person
        WHERE
          cb_search_person.person_id IN (
            SELECT
              criteria.person_id
            FROM
              (SELECT
                DISTINCT person_id,
                entry_date,
                concept_id
              FROM
                `cb_search_all_events`
              WHERE
                (
                  concept_id IN (
                    SELECT
                      DISTINCT c.concept_id
                    FROM
                      `cb_criteria` c
                    JOIN
                      (
                        select
                          cast(cr.id as string) as id
                        FROM
                          `cb_criteria` cr
                        WHERE
                          concept_id IN (4299535)
                          AND full_text LIKE '%_rank1]%'
```

```
                        ) a
                           ON (
                              c.path LIKE CONCAT('%.',
                           a.id,
                           '.%')
                           OR c.path LIKE CONCAT('%.',
                           a.id)
                           OR c.path LIKE CONCAT(a.id,
                           '.%')
                           OR c.path = a.id)
                        WHERE
                           is_standard = 1
                           AND is_selectable = 1
                        )
                        AND is_standard = 1
                     )
                  ) criteria
               ) ))", sep="")

survey_68612580_path <- file.path(
  Sys.getenv("WORKSPACE_BUCKET"),
  "bq_exports",
  Sys.getenv("OWNER_EMAIL"),
  "survey_68612580",
  "survey_68612580_*.csv")
message(str_glue('The data will be written to {survey_68612580_path}. Use this path when
reading ',
                 'the data into your notebooks in the future.'))

bq_table_save(
  bq_dataset_query(Sys.getenv("WORKSPACE_CDR"), dataset_68612580_survey_sql, billing =
Sys.getenv("GOOGLE_PROJECT")),
  survey_68612580_path,
  destination_format = "CSV")

read_bq_export_from_workspace_bucket <- function(export_path) {
  col_types <- NULL
  bind_rows(
    map(system2('gsutil', args = c('ls', export_path), stdout = TRUE, stderr = TRUE),
        function(csv) {
          message(str_glue('Loading {csv}.'))
          chunk <- read_csv(pipe(str_glue('gsutil cat {csv}')), col_types = col_types,
show_col_types = FALSE)
          if (is.null(col_types)) {
            col_types <- spec(chunk)
```

```
        }
        chunk
      }))
}
dataset_68612580_survey_df <-
read_bq_export_from_workspace_bucket(survey_68612580_path)

library(tidyverse)
library(bigrquery)

# This query represents dataset "DataSet_10_13" for domain "person" and was generated for All
of Us Registered Tier Dataset v5
dataset_68612580_person_sql <- paste("
   SELECT
      person.person_id,
      person.gender_concept_id,
      p_gender_concept.concept_name as gender,
      person.birth_datetime as date_of_birth,
      person.race_concept_id,
      p_race_concept.concept_name as race,
      person.ethnicity_concept_id,
      p_ethnicity_concept.concept_name as ethnicity,
      person.sex_at_birth_concept_id,
      p_sex_at_birth_concept.concept_name as sex_at_birth
   FROM
      `person` person
   LEFT JOIN
      `concept` p_gender_concept
         ON person.gender_concept_id = p_gender_concept.concept_id
   LEFT JOIN
      `concept` p_race_concept
         ON person.race_concept_id = p_race_concept.concept_id
   LEFT JOIN
      `concept` p_ethnicity_concept
         ON person.ethnicity_concept_id = p_ethnicity_concept.concept_id
   LEFT JOIN
      `concept` p_sex_at_birth_concept
         ON person.sex_at_birth_concept_id = p_sex_at_birth_concept.concept_id
   WHERE
      person.PERSON_ID IN (
         SELECT
            distinct person_id
         FROM
            `cb_search_person` cb_search_person
         WHERE
```

```
cb_search_person.person_id IN (
   SELECT
      criteria.person_id
   FROM
      (SELECT
         DISTINCT person_id,
         entry_date,
         concept_id
      FROM
         `cb_search_all_events`
      WHERE
         (
            concept_id IN (
               SELECT
                  DISTINCT c.concept_id
               FROM
                  `cb_criteria` c
               JOIN
                  (
                     select
                        cast(cr.id as string) as id
                     FROM
                        `cb_criteria` cr
                     WHERE
                        concept_id IN (4299535)
                        AND full_text LIKE '%_rank1]%'
                  ) a
                     ON (
                        c.path LIKE CONCAT('%.',
                     a.id,
                     '.%')
                     OR c.path LIKE CONCAT('%.',
                     a.id)
                     OR c.path LIKE CONCAT(a.id,
                     '.%')
                     OR c.path = a.id)
               WHERE
                  is_standard = 1
                  AND is_selectable = 1
               )
               AND is_standard = 1
         )
      ) criteria
   ) )", sep="")
```

```
person_68612580_path <- file.path(
  Sys.getenv("WORKSPACE_BUCKET"),
  "bq_exports",
  Sys.getenv("OWNER_EMAIL"),
  "person_68612580",
  "person_68612580_*.csv")
message(str_glue('The data will be written to {person_68612580_path}. Use this path when
reading ',
                 'the data into your notebooks in the future.'))

bq_table_save(
  bq_dataset_query(Sys.getenv("WORKSPACE_CDR"), dataset_68612580_person_sql, billing =
Sys.getenv("GOOGLE_PROJECT")),
  person_68612580_path,
  destination_format = "CSV")

read_bq_export_from_workspace_bucket <- function(export_path) {
  col_types <- NULL
  bind_rows(
    map(system2('gsutil', args = c('ls', export_path), stdout = TRUE, stderr = TRUE),
        function(csv) {
          message(str_glue('Loading {csv}.'))
          chunk <- read_csv(pipe(str_glue('gsutil cat {csv}')), col_types = col_types,
show_col_types = FALSE)
          if (is.null(col_types)) {
            col_types <- spec(chunk)
          }
          chunk
        }))
}
dataset_68612580_person_df <-
read_bq_export_from_workspace_bucket(person_68612580_path)
```

Code to modify survey responses:

```
surv_dat = dataset_68612580_survey_df

# # re-labeling questions

q_label = rep("empty", nrow(surv_dat))

for (i in 1:nrow(surv_dat)){
   if (surv_dat$question_concept_id[i]==43530439){
      q_label[i]="Util4"
```

```
   } else if (surv_dat$question_concept_id[i]==43530437){
      q_label[i]="Util2"
   } else if (surv_dat$question_concept_id[i]==43530438){
      q_label[i]="Util3"
   } else if (surv_dat$question_concept_id[i]==43530401){
      q_label[i]="UtilX"
   } else if (surv_dat$question_concept_id[i]==1585766){
      q_label[i]="Overall_1"
   } else if (surv_dat$question_concept_id[i]==1585778){
      q_label[i]="Overall_3"
   } else if (surv_dat$question_concept_id[i]==1585772){
      q_label[i]="Overall_2"
   } else {q_label[i]="Other"}
}

# re-labeling & merging answers

am_label=c()
am_value=rep(0, nrow(surv_dat))

for (i in 1:nrow(surv_dat)){
   if (surv_dat$answer_concept_id[i]==43529553){
      am_label[i]="Yes"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==43530247){
      am_label[i]="No"
      am_value[i]=-1
   } else if (surv_dat$answer_concept_id[i]==903096){
      am_label[i]="Skip"
      am_value[i]=NA
   } else if (surv_dat$answer_concept_id[i]==903087){
      am_label[i]="DontKnow"
      am_value[i]=NA
   } else if (surv_dat$answer_concept_id[i]==0){
      am_label[i]="DidntAnswer"
      am_value[i]=NA
   } else if (surv_dat$answer_concept_id[i]==1585767){
      am_label[i]="Extremely_Quite"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==1585768){
      am_label[i]="Extremely_Quite"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==1585769){
      am_label[i]="Somewhat_Alittle"
      am_value[i]=0
```

```
} else if (surv_dat$answer_concept_id[i]==1585770){
   am_label[i]="Somewhat_Alittle"
   am_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585771){
   am_label[i]="NotAtAll"
   am_value[i]=-1
} else if (surv_dat$answer_concept_id[i]==1585773){
   am_label[i]="Always_Often"
   am_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585774){
   am_label[i]="Always_Often"
   am_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585775){
   am_label[i]="Sometimes_Occasionally"
   am_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585776){
   am_label[i]="Sometimes_Occasionally"
   am_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585777){
   am_label[i]="Never"
   am_value[i]=-1
} else if (surv_dat$answer_concept_id[i]==1585779){
   am_label[i]="Always_Often"
   am_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585780){
   am_label[i]="Always_Often"
   am_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585781){
   am_label[i]="Sometimes_Occasionally"
   am_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585782){
   am_label[i]="Sometimes_Occasionally"
   am_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585783){
   am_label[i]="Never"
   am_value[i]=-1
} else if (surv_dat$answer_concept_id[i]==43528386){
   am_label[i]="Always_Most"
   am_value[i]=1
} else if (surv_dat$answer_concept_id[i]==43529238){
   am_label[i]="Always_Most"
   am_value[i]=1
} else if (surv_dat$answer_concept_id[i]==43529843){
   am_label[i]="Some"
   am_value[i]=0
```

```
   } else if (surv_dat$answer_concept_id[i]==43529575){
      am_label[i]="None"
      am_value[i]=-1
   } else if (surv_dat$answer_concept_id[i]==43528388){
      am_label[i]="Always_Most"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==43529240){
      am_label[i]="Always_Most"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==43529845){
      am_label[i]="Some"
      am_value[i]=0
   } else if (surv_dat$answer_concept_id[i]==43529577){
      am_label[i]="None"
      am_value[i]=-1
   } else if (surv_dat$answer_concept_id[i]==43528390){
      am_label[i]="Always_Most"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==43529242){
      am_label[i]="Always_Most"
      am_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==43529847){
      am_label[i]="Some"
      am_value[i]=0
   } else if (surv_dat$answer_concept_id[i]==43529579){
      am_label[i]="None"
      am_value[i]=-1
   } else {
      am_label[i]="Other"
      am_value[i]=NA}
}


# re-labeling & NOT merging answers

a_label=c()
a_value=rep(0, nrow(surv_dat))

for (i in 1:nrow(surv_dat)){
   if (surv_dat$answer_concept_id[i]==43529553){
      a_label[i]="Yes"
      a_value[i]=1
   } else if (surv_dat$answer_concept_id[i]==43530247){
      a_label[i]="No"
      a_value[i]=-1
```

```
} else if (surv_dat$answer_concept_id[i]==903096){
   a_label[i]="Skip"
   a_value[i]=0
} else if (surv_dat$answer_concept_id[i]==903087){
   a_label[i]="DontKnow"
   a_value[i]=NA
} else if (surv_dat$answer_concept_id[i]==0){
   a_label[i]="DidntAnswer"
   a_value[i]=NA
} else if (surv_dat$answer_concept_id[i]==1585767){
   a_label[i]="Extremely"
   a_value[i]=4
} else if (surv_dat$answer_concept_id[i]==1585768){
   a_label[i]="Quite"
   a_value[i]=3
} else if (surv_dat$answer_concept_id[i]==1585769){
   a_label[i]="Somewhat"
   a_value[i]=2
} else if (surv_dat$answer_concept_id[i]==1585770){
   a_label[i]="Alittle"
   a_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585771){
   a_label[i]="NotAtAll"
   a_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585773){
   a_label[i]="Always"
   a_value[i]=4
} else if (surv_dat$answer_concept_id[i]==1585774){
   a_label[i]="Often"
   a_value[i]=3
} else if (surv_dat$answer_concept_id[i]==1585775){
   a_label[i]="Sometimes"
   a_value[i]=2
} else if (surv_dat$answer_concept_id[i]==1585776){
   a_label[i]="Occasionally"
   a_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585777){
   a_label[i]="Never"
   a_value[i]=0
} else if (surv_dat$answer_concept_id[i]==1585779){
   a_label[i]="Always"
   a_value[i]=4
} else if (surv_dat$answer_concept_id[i]==1585780){
   a_label[i]="Often"
   a_value[i]=3
```

```
} else if (surv_dat$answer_concept_id[i]==1585781){
   a_label[i]="Sometimes"
   a_value[i]=2
} else if (surv_dat$answer_concept_id[i]==1585782){
   a_label[i]="Occasionally"
   a_value[i]=1
} else if (surv_dat$answer_concept_id[i]==1585783){
   a_label[i]="Never"
   a_value[i]=0
} else if (surv_dat$answer_concept_id[i]==43528386){
   a_label[i]="Always"
   a_value[i]=4
} else if (surv_dat$answer_concept_id[i]==43529238){
   a_label[i]="Most"
   a_value[i]=3
} else if (surv_dat$answer_concept_id[i]==43529843){
   a_label[i]="Some"
   a_value[i]=2
} else if (surv_dat$answer_concept_id[i]==43529575){
   a_label[i]="None"
   a_value[i]=1
} else if (surv_dat$answer_concept_id[i]==43528388){
   a_label[i]="Always"
   a_value[i]=4
} else if (surv_dat$answer_concept_id[i]==43529240){
   a_label[i]="Most"
   a_value[i]=3
} else if (surv_dat$answer_concept_id[i]==43529845){
   a_label[i]="Some"
   a_value[i]=2
} else if (surv_dat$answer_concept_id[i]==43529577){
   a_label[i]="None"
   a_value[i]=1
} else if (surv_dat$answer_concept_id[i]==43528390){
   a_label[i]="Always"
   a_value[i]=4
} else if (surv_dat$answer_concept_id[i]==43529242){
   a_label[i]="Most"
   a_value[i]=3
} else if (surv_dat$answer_concept_id[i]==43529847){
   a_label[i]="Some"
   a_value[i]=2
} else if (surv_dat$answer_concept_id[i]==43529579){
   a_label[i]="None"
   a_value[i]=1
```

```
   } else {
      a_label[i]="Other"
      a_value[i]=NA}
}

# # creating merged Q&A's

surv_dat_qa = cbind(surv_dat, q_label, a_label, a_value,am_value)
surv_dat_qa[1:3,]
```

Code to separate GDM specific data:

```
### Simplify dataset names

cond_dat = dataset_68612580_condition_df
demo_dat = dataset_68612580_person_df
patients = unique(cond_dat$person_id)


# flag patient IDs with and without GDM

# 4024659 -- Gestational Diabetes mellitus
# 4263902 -- Gestational diabetes mellitus, class A>2<
# 4326434 -- Gestational diabetes mellitus, class A>1<

gdm_id = c(4024659, 4263902, 4326434)

ind.gdm = which(cond_dat$condition_concept_id %in% gdm_id)

patient_gdm = unique(cond_dat$person_id[ind.gdm])
patient_no_gdm = setdiff(patients, patient_gdm)

   # surveys for GDM and nonGDM patients

survey_gdm = surv_dat_qa %>% filter(person_id %in% patient_gdm)
survey_no_gdm = surv_dat_qa %>% filter(person_id %in% patient_no_gdm)

# create gdm flag

gdm_flag=rep("empty", nrow(surv_dat_qa))

for (i in 1:nrow(surv_dat_qa)){
   if (surv_dat_qa$person_id[i] %in% patient_gdm){
```

```
      gdm_flag[i]="GDM"} else {gdm_flag[i]="no GDM"}
   }

gdm_flag = as.factor(gdm_flag)

surv_dat_qa_gdm = cbind(surv_dat_qa, gdm_flag)

   # demographic data for GDM and nonGDM patients

demo_gdm = demo_dat %>% filter(person_id %in% patient_gdm)
demo_no_gdm = demo_dat %>% filter(person_id %in% patient_no_gdm)

   # all conditions for GDM and nonGDM patients

cond_gdm = cond_dat %>% filter(person_id %in% patient_gdm)
cond_no_gdm = cond_dat %>% filter(person_id %in% patient_no_gdm)
```

Code to test association between survey responses and GDM status:

```
survey = surv_dat_qa_gdm

# # Mann-Whitney U Tests

#### Utilization Q2

ind.u2 = which(survey$q_label=="Util2")
data.u2 = survey[ind.u2,10:14]

#str(data.u2)
#nrow(data.u2) - sum(is.na(data.o1$a_value))

# merged
wilcox.test(am_value ~ gdm_flag, data.u2)
ggplot() + geom_density(aes(x = am_value, fill = gdm_flag), alpha = 0.5, data=data.u2)

# not merged
wilcox.test(a_value ~ gdm_flag, data.u2)
ggplot() + geom_density(aes(x = a_value, fill = gdm_flag), alpha = 0.5, data=data.u2)

#### Utilization Q3

ind.u3 = which(survey$q_label=="Util3")
data.u3 = survey[ind.u3,10:14]

#str(data.u2)
```

```
#nrow(data.u2) - sum(is.na(data.o1$a_value))

# merged
wilcox.test(am_value ~ gdm_flag, data.u3)
ggplot() + geom_density(aes(x = am_value, fill = gdm_flag), alpha = 0.5, data=data.u3)

# not merged
wilcox.test(a_value ~ gdm_flag, data.u3)
ggplot() + geom_density(aes(x = a_value, fill = gdm_flag), alpha = 0.5, data=data.u3)

#### Utilization Q4

ind.u4 = which(survey$q_label=="Util4")
data.u3 = survey[ind.u4,10:14]

#str(data.u2)
#nrow(data.u2) - sum(is.na(data.o1$a_value))

# merged
wilcox.test(am_value ~ gdm_flag, data.u3)
ggplot() + geom_density(aes(x = am_value, fill = gdm_flag), alpha = 0.5, data=data.u3)

# not merged
wilcox.test(a_value ~ gdm_flag, data.u3)
ggplot() + geom_density(aes(x = a_value, fill = gdm_flag), alpha = 0.5, data=data.u3)

#### Overall Q1

ind.o1 = which(survey$q_label=="Overall_1")
data.o1 = survey[ind.o1,10:14]

str(data.o1)
nrow(data.o1) - sum(is.na(data.o1$a_value))

# merged
wilcox.test(am_value ~ gdm_flag, data.o1)
ggplot() + geom_density(aes(x = am_value, fill = gdm_flag), alpha = 0.5, data=data.o1)

# not merged
wilcox.test(a_value ~ gdm_flag, data.o1)
ggplot() + geom_density(aes(x = a_value, fill = gdm_flag), alpha = 0.5, data=data.o1)

#### Overall Q2

ind.o2 = which(survey$q_label=="Overall_2")
```

```
data.o2 = survey[ind.o2,10:14]

str(data.o2)
nrow(data.o2) - sum(is.na(data.o2$a_value))

# merged
wilcox.test(am_value ~ gdm_flag, data.o2)
ggplot() + geom_density(aes(x = am_value, fill = gdm_flag), alpha = 0.5, data=data.o2)

# not merged
wilcox.test(a_value ~ gdm_flag, data.o2)
ggplot() + geom_density(aes(x = a_value, fill = gdm_flag), alpha = 0.5, data=data.o2)

#### Overall Q3

ind.o3 = which(survey$q_label=="Overall_3")
data.o3 = survey[ind.o3,10:14]

str(data.o3)
nrow(data.o3) - sum(is.na(data.o3$a_value))

# merged
wilcox.test(am_value ~ gdm_flag, data.o3)
ggplot() + geom_density(aes(x = am_value, fill = gdm_flag), alpha = 0.5, data=data.o3)

# not merged
wilcox.test(a_value ~ gdm_flag, data.o3)
ggplot() + geom_density(aes(x = a_value, fill = gdm_flag), alpha = 0.5, data=data.o3)
```

Code to flag anxiety & depression data:

```
###### person ids with anxiety

# 434613 -- Generalized anxiety disorder
# 442077 -- Anxiety disorder

anx_id = c(434613, 442077)
ind.anx = which(cond_dat$condition_concept_id %in% anx_id)
patient_anx = unique(cond_dat$person_id[ind.anx])
patient_no_anx = setdiff(patients, patient_anx)


# data for gdm patients with anxiety
survey_gdm_anx = survey_gdm %>% filter(person_id %in% patient_anx)
```

```
demo_gdm_anx = demo_gdm %>% filter(person_id %in% patient_anx)

# # # gdm and anxiety

ind.gdm_anx = which(cond_gdm$condition_concept_id %in% anx_id)
patient_gdm_anx = unique(cond_gdm$person_id[ind.gdm_anx])
patient_gdm_no_anx = setdiff(patient_gdm, patient_gdm_anx)

# # # NO gdm and anxiety

ind.no_gdm_anx = which(cond_no_gdm$condition_concept_id %in% anx_id)
patient_no_gdm_anx = unique(cond_no_gdm$person_id[ind.no_gdm_anx])
patient_no_gdm_no_anx = setdiff(patient_no_gdm, patient_no_gdm_anx)

###### person ids with depression

# 4077577 -- Major recurrent major depression
# 4152280 -- Major depressive disorder

dep_id = c(4077577, 4152280)
ind.dep = which(cond_dat$condition_concept_id %in% dep_id)
patient_dep = unique(cond_dat$person_id[ind.dep])
patient_no_dep = setdiff(patients, patient_dep)

# data for gdm patients with depression
survey_gdm_dep = survey_gdm %>% filter(person_id %in% patient_dep)
demo_gdm_dep = demo_gdm %>% filter(person_id %in% patient_dep)

# # # gdm and depression

ind.gdm_dep = which(cond_gdm$condition_concept_id %in% dep_id)
patient_gdm_dep = unique(cond_gdm$person_id[ind.gdm_dep])
patient_gdm_no_dep = setdiff(patient_gdm, patient_gdm_dep)

# # # NO gdm and depression

ind.no_gdm_dep = which(cond_no_gdm$condition_concept_id %in% dep_id)
patient_no_gdm_dep = unique(cond_no_gdm$person_id[ind.no_gdm_dep])
patient_no_gdm_no_dep = setdiff(patient_no_gdm, patient_no_gdm_dep)
```

Code to perform independence tests for anxiety/depression & GDM:

```
#survey_gdm_dep[1:10,]
```

```
length(patients) #7069

# gdm counts
length(patient_gdm) #1037
length(patient_no_gdm) #6032

# anxiety counts
length(patient_anx) #4244
length(patient_no_anx) #2825

# depression counts
length(patient_dep) #787
length(patient_no_dep) #6282

# gdm & anxiety counts
length(patient_gdm_anx) #308
length(patient_gdm_no_anx) #729

# gdm & depression counts
length(patient_gdm_dep) #55
length(patient_gdm_no_dep) #982

# NO gdm & anxiety counts
length(patient_no_gdm_anx) #3936
length(patient_no_gdm_no_anx) #2096

# NO gdm & depression counts
length(patient_no_gdm_dep) #732
length(patient_no_gdm_no_dep) #5300

### Chi-Square test of independence of gdm & anxiety

# table of counts

# gdm & anxiety counts
gdm_anx=308
gdm_no_anx=729
# NO gdm & anxiety counts
no_gdm_anx=3936
no_gdm_no_anx=2096


tab_gdm_anx <- as.table(rbind(c(gdm_anx,gdm_no_anx), c(no_gdm_anx,no_gdm_no_anx)))
dimnames(tab_gdm_anx) <- list(GDM = c("Y", "N"), Anxiety = c("Y", "N"))
tab_gdm_anx
```

```
chisq.test(tab_gdm_anx)


### Chi-Square test of independence of gdm & depression

# table of counts

# gdm & depression counts
gdm_dep=55
gdm_no_dep=982
# NO gdm & depression counts
no_gdm_dep=732
no_gdm_no_dep=5300

tab_gdm_dep <- as.table(rbind(c(gdm_dep,gdm_no_dep), c(no_gdm_dep,no_gdm_no_dep)))
dimnames(tab_gdm_dep) <- list(GDM = c("Y", "N"), Depression = c("Y", "N"))
tab_gdm_dep

chisq.test(tab_gdm_dep)
```

Code to generate demographic plots:

```
GDM_status = c(rep("GDM",nrow(demo_gdm)), rep("no GDM", nrow(demo_no_gdm)))
Gender = c(demo_gdm$gender, demo_no_gdm$gender)
Race = c(demo_gdm$race, demo_no_gdm$race)
Ethnicity = c(demo_gdm$ethnicity, demo_no_gdm$ethnicity)

data_demos = data.frame(GDM_status, Gender, Race, Ethnicity)
table(data_demos[,1:2])
table(data_demos[,c(1,3)])
table(data_demos[,c(1,4)])

options(repr.plot.width=10, repr.plot.height=6)

ggplot(data_demos, aes(GDM_status,fill=Gender)) +
   geom_bar(position="dodge")

ggplot(data_demos, aes(GDM_status,fill=Race)) +
   geom_bar(position="dodge")

ggplot(data_demos, aes(GDM_status,fill=Ethnicity)) +
   geom_bar(position="dodge")
```