**Gestational Diabetes**

**Objectives:**

1. Identify risk factors for Gestational Diabetes.
2. To understand the impact of GD on maternal and fetal health.
3. Does race and ethnicity have an impact on GD?
4. Do women with high BP before or during pregnancy are at a higher risk of developing gestational diabetes?
5. Does Age and BMI play a role in developing GD?
6. Do women who are diagnosed with gestational diabetes later in their pregnancy are at a higher risk of complications?
7. Do Women who have had gestational diabetes in a previous pregnancy more likely to develop it again in future pregnancies?
8. Is there an association between low hemoglobin levels (anemia) and an increased risk of gestational diabetes?
9. Does uncontrolled GD affect kidneys?
10. What are the complications that can occur for women with GD and their rate of occurrence?
11. Does Vit D play a major role in gestational diabetes?

**Data Pre-processing:**

We used **Python** for the initial cleaning process. We then uploaded the modified file to Excel and used **Power Query Editor** in Excel to finish further cleaning steps.

Read the data into a data frame: Rows – 600 / Columns - 259

**Cleaning Steps:**

1. **Removed the duplicate columns in the data** - 27 columns were removed.

*Dropped columns: Date of Data Collection.1, Miscarriage, V1 Creatinine.1, Date of*

*Delivery, Screened 10 3, Miscarriage before 28/40 2, Perinatal death 10, HELLP 10, GCT 10, Nutrition 10, Screened 10, V3 Creat.1, GCT 1h value >7.5mM 10.1, GCT Ref to DNS 10.1, Lost to f/u, Miscarriage.1, systolic BP (mmHg) V3.1, Miscarriage before 28/40 v2, V3 GA, Fetal hypoglycemia 10, GDM screening performed?, Instrumentation 10, Screening DNA 2,*

*BMI (kg/m2) V1 2, 25OHD 10, Withdrew before 28/40 2*

1. **Removed columns with more than 60% of null values** – 13 columns were removed.

*Dropped columns: ['Medications V1', 'V1 25OHD value (nmol/L)', 'Date 25OHD result*

*received', 'Prescription for Caltrate sent?', 'Date Prescription sent?', 'OGTT 10 V3', 'OGTT 2h value', 'OGTT 1h value high 10.1', 'OGTT Ref to DNS', ‘Cause of fetal death', 'Took Vit D*

*Supplements', 'Insulin/Metformin', 'Twins weight']*

1. **Irrelevant columns** w.r.t. Weight, BMI, Height, HbA1c, Hb, Date, Lab values, and other columns **were removed**. After this process, there are 121 columns and 600 rows.

*Dropped Columns: ['Screened 10 2', 'Smoking 123',' Smoking 10 V1', 'White 10 V1',' Age >30 10',' Obese 10','Overweight 10',' List used',' Vit D List used', '1st DASS score >33',' Ref to Psych',' V1 HBa1C <48 mmol/mol',' Ref to DNS 10',' Withdrew after 28/40 10',' PreviousGDM10 V1','GCT 1h value >7.5mM','3rd visit DASS >33',' AC >350mm 10', 'HighRisk 10','25OHD checked yes (1) or no (0)', 'GCT 1h value >7.2mM 10',*

*'Number obstetric clinics',' Glucose lowering therapy 10',' GHP 10',' Pre-eclampsia 10',' OGTT 0h value',' OGTT 0h value high 10', 'OGTT 1h value', 'OGTT 1h value high 10', 'OGTT values high',' OGTT high 10',' Eclampsia 10', 'HELLP',' Induction 10',' LSCS 10', 'GCT 1h value >7.5mM 10', 'GCT 1h value >7.8mM 10', 'Still-birth 10',' Birth weight >4kg',' BW >4kg 10',' BW >4.5kg 10',' BW <2.5kg 10',' OGTT performed',' Birth Injury 10',' Dystocia 10', 'Brachial plexus injury',' Plexus injury 10',' Other nerve injury', ' Clavicular fracture', 'Humeral fracture', 'Skull fracture’, ‘Other birth injury',' Perinatal death’, ‘CT Ref to DNS 10',' Fetal jaundice 10',' Fetal phototherapy 10',' RR>60',' RR>60 10',' Cong malformation 10',' Spina bifida', 'Cleft palate',' Clef Lip', 'Hydrocephalus',' Cardiac murmur',' Cardiac anomaly', 'Tetralogy of Fallot', 'Transposition of Great arteries', 'Septal defect',' Other cardiac anomaly', 'OGTT Ref to DNS 10',' DNA Visit 3 10', 'Attended Visit 3 10',' Ref to psych',' V3 AC <35 cm', 'm Weight (kg) V1',' systolic BP (mmHg) V1.1',' V1 Hb.1',' V1 ALT.1','V1 CRP.1',' V1 PCR.1',' Attended GDM clinics',' Number GDM clinics',' Nutritional counselling',' Surfactant use', 'SCBU 10']*

1. **Structural errors**- ‘Date of visit 3’ and ‘Date of Data Collection’ columns had the date and text values in the column. We extracted ‘date’ and ‘text’ into two different columns.
2. **Handling missing values**: The lab values column in the dataset with null values was replaced with **the mean value of the respective column**.

*Imputed columns: ['V1 HbA1c (mmol/mol)', 'WCC',' V1 Hb',' V1 Platelet',' V1 Creatinine',' V1 ALT',' V1 CRP', 'V1 U protein', 'V1 U creatinine',' V1 PCR',' systolic BP (mmHg) V1',' diastolic BP (mmHg) V1']*

Further cleaning was carried out in **Power Query Editor** in Excel.

1. The *Gestational Age* column was in the format weeks+ days, changed the format into decimal.
2. *BMI, V1 WCC, V1 Hb,V1 Platelet, V! ALT, V1Creatinine, V1 Protein,* column was rounded to 1 decimal points.
3. ***Changed datatypes****: EDD* column was changed to Date format*, V3 25OHD* changed to Decimal number and rounded to 2 digis.
4. The **null values** in the *columns Smoking, GCT, 25OHD Checked, GCT OGTT high 10, Diagnosed with Vit D deficiency, Dx with GDM , Eclampsia, Emergency, Birth Injury* **were replaced with the string ‘Not Recorded’.**
5. Replaced Yes-1, No-0: Previous GDM, Age>30, Hx GDM, glycosuria, 1st degree rel
6. **Renaming Column Names** – *Overweight -BMI Category, Hx GDM, glycosuria, 1st degree rel – High Risk*
7. **Deleted columns**: *US EDD, Screening DNA, Abdominal Circumference*
8. **Replaced null with 0** - *1 hr Glucose. Weight(kg), Systolic BP (mmHg) V3, diastolic BP (mmHg) V3, V3 HbA1c, V3 WCC, V3, Hb, V3, Platelet, V3 U Creat, V3 ALT, V3 U PCR, 25 OHD value, m V3 Weight, Weight Change BBP change, BP change (%), HbA1c change, Hb change, Create change(%), ALT change, CRP change, GA at delivery.*
9. **Inaccurate entries or data entry errors** such as ‘NO’, and ’N0’ were replaced with ‘No’, ‘yes’ was replaced with ‘Yes’, and the null values were replaced with ‘Not recorded’ for the following columns: *GCT, Dx with GDM, Glucose-Lowering Therapies, Gestational Hypertension, Preeclampsia, Induction, Caesarian, Instrumentation, Still Birth, Shoulder Dystopia, Antenatal Steroid Use, Fetal Hypoglycemia, Fetal Jaundice, Fetal Phototherapy, Total Bilirubin, SCBU, Cong Malformation, Epidural, Spinal*

**Data Definition:**

Our cleaned dataset now consists of 600 observations and 86 variables.

|  |  |  |  |
| --- | --- | --- | --- |
| S.NO | **Biomarkers** | Explanation | Ranges |
| 1 | **Participant ID** | Unique ID | - |
| 2 | **GCT** | The glucose challenge test, also called the one-hour glucose tolerance test, measures your body's response to sugar (glucose). | 0/1 |
| 3 | **Systolic BP (mmHg)** | systolic blood pressure, measures the pressure in your arteries when your heart beats. | below 120 mmHg |
| 4 | **Diastolic BP (mmHg)** | diastolic blood pressure, measures the pressure in your arteries when your heart rests between beats. | below 80 mmHg |
| 5 | **Weight (kg)** | Weight of body | weight gain  Normal weight (BMI 18.5-24.9): 25-35 pounds (11.3-15.9 kg) |
| 6 | **BMI (kg/m2)** | Body mass index (BMI) is a measure of body fat based on height and weight that applies to adult men and women. | BMI 18.5-24.9 |
| 7 | **BMI category** | Normal, obese, overweight |  |
| 8 | **Smoking** | Tobacco Intake | Yes/No |
| 9 | **Alcohol Intake** | Habit of Alcohol intake | Yes/No |
| 10 | **Ethnicity** | An ethnicity or ethnic group is a grouping of people who identify with each other on the basis of perceived shared attributes that distinguish them from other groups. | Asian, Black, White, Mixed Race |
| 11 | **Previous GDM** | Previous history of Gestational Diabetes Miletus | 0/1 |
| 12 | **Chronic Illness** | causes insulin resistance, which is a key factor in the development of gestational diabetes. | 0/1 |
| 13 | **Meds 10** | Medications were prescribed or not | 0/1 |
| 14 | **Age>30** | Age less than 30 having GDM or not | - |
| 15 | **High Risk GDM** | Patient is high risk with GDM or not | 0/1 |
| 16 | **Screening method** | OGTT or GCT | - |
| 17 | **25OHD checked** | 25-hydroxyvitamin D (25(OH)D), this is a test that measures the level of vitamin D in the blood | Yes/no |
| 18 | **Gestational Age** | age of the fetus or pregnancy | no. of weeks |
| 19 | **EDD** | Expected Delivery Date | - |
| 20 | **HbA1c (mmol/mol)** | (hemoglobin A1c) is a blood test used to measure the average blood sugar (glucose) | 42 mmol/mol-for non-pregnant women  Less than 48 mmol/mol |
| 21 | **WCC** | white cell count, high cell count lead to inflammation, low cell count ( bone marrow suppression ) | 4.0 and 11.0 |
| 22 | **Hb** | hemoglobin(is used to help diagnose and monitor diabetes) | 11 and 12.5 |
| 23 | **Platelet** | gestational thrombocytopenia, a condition in which platelet levels drop below normal levels can increase the risk of bleeding and bruising during childbirth | 150-450 |
| 24 | **Creatinine** | gestational diabetes can affect kidney function and lead to a buildup of waste products in the blood, including creatinine. | 53 μmol/l (0.60 mg/dl),10 |
| 25 | **ALT** | alanine transaminase, cause liver damage | 7 and 56 units per liter (U/L) |
| 26 | **U albumin** | Urine albumin level | less than 30 milligrams per gram of creatinine |
| 27 | **U protein** | Urine protein level | 6.0 and 8.3(g/dL). |
| 28 | **U creatinine** | Urine creatinine level | 2.66-28.38mmol/L |
| 29 | **U PCR** | Urine protein and creatinine ratio | less than 0.3 milligrams than creatinine |
| 30 | **CRP** | C-reactive protein, Elevated CRP led to inflammation increased risk of complications such as pre-eclampsia and premature labor | less than 10(mg/L) |
| 31 | **GDM screening performed? V2** | gestational diabetes mellitus (24 and 28 weeks) risk factors-history of GDM | Y/N/Transferred |
| 32 | **Miscarriage before 28/40** | Miscarried before 28 weeks | Y/N |
| 33 | **Withdrew before 28/40** | withdrawing from the treatment before28 weeks | Y/N |
| 34 | **Screening DNA** | Suggested for DNA test | Y/N |
| 35 | **1h glucose** | One hour glucose test | 7.8 mmol/L |
| 36 | **Miscarriage after 28/40** | Miscarried after 28 weeks | Y/N |
| 37 | **Delivered before 36/40** | Delivered early | Y/N |
| 38 | **Dx with GDM** | Diagnosed with GDM | Y/N |
| 39 | **Diagnosed with Vit D Deficiency** | Vitamin D deficiency diagnosed or not | Y/N |
| 40 | **Insulin** | Insulin therapy is used for women with GDM who are unable to control their blood sugar levels through diet and exercise | 0/1 |
| 41 | **Metformin** | Metformin is an oral medication used to treat type 2 diabetes by lowering blood sugar levels | 0/1 |
| 42 | **Gestational hypertension** | high blood pressure during pregnancy, typically after 20 weeks of gestation | Y/N |
| 43 | **Pre-eclampsia** | Is a serious complication of pregnancy, characterized by high BP and damage to organs such as the liver and kidneys | Y/N |
| 44 | **Eclampsia** | Is a serious complication of pre-eclampsia | Y/N |
| 45 | **Induction** | Inducing labor | Y/N |
| 46 | **Caesarean** | Operated | Y/N |
| 47 | **Emergency** | Taken to ER | Y/N |
| 48 | **Instrumentation** | forceps or a ventouse suction cup are used to help deliver the baby | Y/N |
| 49 | **Still-birth** | A stillbirth is the death or loss of a baby before or during delivery. | Y/N |
| 50 | **Twins** | Delivered twins or not | Y/N |
| 51 | **Birth weight (kg)** | During birth, weight of the baby | 2.5 and 4.0 kg |
| 52 | **Apgar 1 minute, Apgar 3 minute** | Appearance, Pulse, Grimace, Activity, and Respiration test, five things are used to check a baby's health | score of 0 to 3 is concerning |
| 53 | **Birth Injury** | Any birth injuries during delivery | Y/N |
| 54 | **Shoulder dystocia** | baby's shoulders get stuck during vaginal delivery. | Y/N |
| 55 | **Antenatal steroid use** | medication administered to pregnant women expecting a preterm birth. | Y/N |
| 56 | **Fetal hypoglycemia** | condition in which the level of glucose in the blood of a developing fetus is lower than normal | Y/N |
| 57 | **Fetal jaundice** | increased risk of developing jaundice due to higher levels of bilirubin in their blood | Y/N |
| 58 | **Fetal phototherapy** | Is a safe and effective treatment for neonatal jaundice | Y/N |
| 59 | **SCBU** | Special Care Baby Unit that provides medical care to sick or premature newborn babies | Y/N |
| 60 | **Cong malformation** | congenital malformations may be caused by genetic factors | Y/N |
| 61 | **Epidural** | anesthesia is commonly used during childbirth to provide pain relief for the mother | Y/N |
| 62 | **Spinal** | spinal anesthesia typically takes effect faster than an epidural, and provides more complete pain relief in the lower body | Y/N |
| 63 | **Visit 3 - Reason** |  |  |
| 64 | **Finalstatus** |  |  |

**\*\* 0- No, 1- YES, \*V1,V2,V3-Visit1,2,3, \***