Chapter 3:
Methodology
Overview:
This chapter concentrates on research methodology of this study. Study population, research and tools design, sampling, data collection and analysis used in this study, have been detailed. Also definition and introducing of the variables were considered in this chapter. Furthermore, statistical test used in this study, has been elaborated.
3.1 study population design steps:
The following points were considered in developing the sample design.

3.1.1. Selection of subjects (This section explained in analysis chapter).
3.1.2. Sample size.

Figure 3.1: variables of study and distracting factors

* A variable that can cause or prevent the outcome of interest, and is not intermediate variable and is associated with the factor under investigation. Unless it is possible to adjust for confounding variables their effects cannot be distinguished from those of factor(s) being studied (Dictionary of epidemiology, Last JM, 2001)
3.1.2. **Sample size:**

An optimum sample size is one which fulfills the requirement of efficiency, representation, reliability and flexibility.

The main target of this study was to compare the characteristics of GDM mothers with those of NGT women with special reference to dietary intake of antioxidants (HORAC and vitamin C) and circulating level of vitamin C, with case-control design.

Considering probability of exposure and controls it was decided to assume the ratio of exposed to non-exposed as 1:3 to demonstrate the possible effect. According to literature the event risk for controls was 5%, size of effect in main variable (plasma vitamin C on glycemic status) were 3 (Cuilin Z. et al. 2004).

Using sample size calculation software, PS Power version 2.1.30, and significance level of 5% (α = 0.05), and power 80% (1-β = 0.80) the sample size determined as 42 and 126 for exposed and non-exposed respectively, identical to 168. However in present study, 200 samples has been collected comprise of 42 cases and 158 controls.

**3.2. Research design steps:**

**3.2.1. Ethical considerations:**

Before starting the main body of research according to ICMR protocols, data was completely confidential; each patient was asked to sign an informed consent (It is provided in appendix). The privacy curing interview and clinical procedures was considered. They were explained why they are selected for this research.

No name was used in the study. Participation to this study was voluntarily. In addition, World Medical Association (WMA) Declaration of Helsinki (1964, 2008) and CIOMS ethical considerations (Council of Organization of Medical Sciences) (2002), and National Institute of Health (NIH) 2011, has been considered. This study was presented and approved by KEM hospital research centre ethical committee.

Also sub-part of HHS regulation (on Belmont report 1979 codified 45 CFR46), for “additional protection for pregnant women, human fetuses and neonates” has been considered in this research (PHRP, 2008).

A Subject Information Sheet (SIS) which explains all parts of study in a simple way to understand for different literacy groups were prepared. (Appendix) SIS and
consent form were prepared and translated to Hindi language and for accuracy back translation was also performed. The forms provided the detail information on the study. Voluntary participation and confidentiality of the data was assured. Consent form of each respondent was obtained before start of the study.

<table>
<thead>
<tr>
<th>Study Design</th>
<th>Case control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data source</td>
<td>Pregnant women of 24-28 weeks of gestation attending ANC of KEM Hospital Research Center and Gupte Hospital both in Pune city between May 2012 to December 2012</td>
</tr>
<tr>
<td>Outcome</td>
<td>Glycemic status by OGTT, baby’s birth weight, gestation age at birth</td>
</tr>
<tr>
<td>Explanatory variable</td>
<td>Maternal [Age, demographic data, obstetric/medical history, present obstetric information, delivery information, ultrasounds reports, anthropometric measurements(Height ,weight, BMI), any tobacco consumption, passive/active smoking, dietary habits, physical activity, level of life stress</td>
</tr>
<tr>
<td>Sample size</td>
<td>42 GDM, 158 NGT collected (42 case, 126 controls according to sample size calculation had been determined)</td>
</tr>
</tbody>
</table>

3.2.2. Research tool (Questionnaire):

Personal information was collected through structured interview. Pretested structured questionnaire was used for interview. The key words were translated into Marathi and Hindi, wherever needed for ease of understanding. A medical graduated person was helping to ask the questions. All risk factors of GDM were included in the questionnaire. The questionnaire content covered all study’s variables.

During each interview, information concerning the demographic background of pregnant mothers comprising; of maternal medical and obstetric history, social and educational background, age, religion, maternal work status, number of household members, total family income, were sought. Birth details were recorded from birth record file that was available in obstetric/neonatal department of the hospital, including infant weight and gender, and any complications like, hypoglycemia*, hyperbilirubinemia*, Respiratory Distress Syndrome (RDS)* admission to NICU. Dietary assessment was...
accomplished by 24 hour diet recall (for week day as well as weekend day if there is any difference) and Semi Food Frequency Questionnaire (SFFQ) for rich sources of vitamin C and High Oxygen Absorbance Capacity Foods (HORAC) and a dietary habit questionnaire.

The questionnaire was constructed in twelve sections,

Section A: Maternal demographic information
Section B: Maternal obstetric and medical history
Section C: Maternal antenatal information and anthropometric measurements
Section D: Biochemical parameters
Section E: Antenatal ultra-sound reports
Section F: Maternal present obstetric information and any complication
Section G: Maternal delivery and post partum information
Section H: Neonatal complications
Section I: Dietary habits
Section J: smoking habits
Section K: Life stress during pregnancy
Section L: Pre-pregnancy physical activity history

Since structured questionnaire was used, interview bias was reduced.

In the following, the sections of questionnaire explained in related to each group of collected variables, also definition of key concepts has been brought anywhere was needed.

3.3 Characteristics of study population:

This section was planned to obtain information on age, maternal education, maternal occupation, total household income, family type, number of family members.

A woman with Bad obstetrical history (BOH) is defined as one with previous poor obstetric outcome that can again adversely influence the future pregnancy and labor (Padubidri and Anand 1996). BOH data from multi-para mothers regarding adverse outcome and previous pregnancies like intra uterine growth retardation (IUGR), polyhydraminosus (more than 2000 ml amniotic fluid or amniotic fluid index (AFI) more than 24 Cm (Dorland’s medical dictionary 2005), macrosomia, pre term delivery (delivery
sooner than 37 weeks of gestation, (Kale SD et al, 2005), GDM, etc, was recorded using structured questionnaire. This part was completed by section B of questionnaire (Maternal obstetric and medical history).

This section was comprised of information on family history of diabetes, previous history of gestational diabetes of the mother, interval between previous and present pregnancy, number of gravid (number of pregnancies), para (or parity number of births after 20 weeks, dead or alive), abortion (loss of pregnancy before 20 weeks of gestation (American pregnancy association), death, live history for pregnancies, previous delivery modes (normal delivery or cesarean). This information was provided by subject’s recall as well as medical history record available in subject’s file in the hospital.

Gestational age was derived from the LMP and the sonographic estimate.

3.4. Nutritional assessment:

3.4.1. Anthropometry:

Women were measured to record their weight and height in duplicate. Mean of the two measurements was used in analysis. Pre-pregnancy weight was recorded based on recall. BMI was calculated as weight (kg) / height (m²).

Body size including height, weight, and body mass index (BMI) was recorded at the time of interview using a plastic stereotype and digital scale respectively. To assure accuracy duplicate measurement was performed. Tanita body analyser model BC, 601 was used to measure the weight of pregnant mother.

Before each weight measurement the scale was calibrated with a 10 kg weight device. The height and weight were measured with subjects in light clothing and without shoes and body weight was recorded to the nearest 0.1 Kg whereas height was measured to the nearest 0.1 cm height using a wall fixed stadiometer (plastic tape). Maternal height and weight were measured in duplicate. Mean of two measurements were used in analysis. Weight at the end of pregnancy was found using same scale. And then weight gain at the time of interview and total weight gain during pregnancy was calculated.

For analysis women were also classified according to WHO expert consultation 2004 (<18.5 low, 18.5-24.99 normal, >25 over weight and >30 obese). Although other categories as thin, BMI less than 19.8, average BMI 19.8-26.0, overweight BMI >26.0,
and very overweight, BMI greater than 29.0 also were available from nutrition during pregnancy, Washington DC, national academy of science, 1990 and Seshiah V, diabetes and pregnancy 2012.

Total gestational weight gain was calculated as the difference between a woman’s pre-pregnancy weight or initial weight and her final recorded weight before delivery.

Pre-pregnancy weight and LMP (last menstrual period) of the subject recorded during interview. Gestational age was determined by ultrasound examination.

3.4.2. Dietary intake:

The conventional 24 hour diet recall method was modified and made more objective by incorporating information on portion sizes. Women were interviewed at average of 26 weeks of gestation to record the consumption of food items in chronological order from morning until dinner time, for week and weekend day if there was any difference. In addition to 24-h diet recall questionnaire, a (Semi Food Frequency Questionnaire) SFFQ was administered to obtain frequency of consumption of Vitamin C rich foods and high oxygen radical absorbance capacity (HORAC) foods, during the preceding 6 month period on a six point scale ranging from “never” to “daily”. The frequency of intake by each subject was calculated and converted to per day frequency. For this purpose daily intake, weekly intake and monthly intake were multiplied by 30, 4 and 1 respectively. Then the result devided by 30 to get daily score. The following scores were considered: zero as never taken, 0-4 <once per week, 4-15 > once a week, 15 as every alternate day. Food groups were including fruit and vegetables for vitamin C rich sources and fruit, vegetables, spices, nuts, beans and miscenlonous for HORAC foods.

Questionnaire included history of intake of supplements which containing vitamin C (iron, and other multi-vitamins), and their amounts per day. Intake of other nutritional supplements like Bournvita, Proteinex, Complan, or any medicine like Shatawari, was also recorded, since all these supplements are extra sources for getting vitamin C. For this purpose, during interview by 24 hour diet recall, the generic name of supplement and daily dosage was asked and vitamin C intake from these supplements was calculated accordingly and the calculated amount summed with other taken rich vitamin C foods by each subjects and reported as total daily vitamin C intake by each subject.
Semi food frequency questionnaire (SFFQ) was modified to avail the data for estimating relevant micronutrient consumption by the study population. Hence questionnaire included locally available rich sources of vitamin C. Previously this method was used successfully (Bhaskarachary, et.al, 2010). Food Frequency Questionnaire (FFQ) is contains of all food groups, whereas in SFFQ an only rich source of selected nutrient is provided.

In present study, locally rich sources of vitamin C were included: Alma (gooseberry), Guava, Orange, Lime, Papaya, Lemon, Pineapple, Custard apple, strawberry and other berries, mango, pomegranate (under fruit category) and capsicum, Coriander leaves, Cabbage, Bitter guard, Cauliflower, fenugreek leaves, tomato, spinach, potato, tamarind pulp, drum stick leaves, turnip, carrot, amaranth, sprouted Bengal gram (under vegetable category).

To detail the above discussion, “Diet history questionnaire” employed in this study encouraged a narrative style of reporting by participants but allowed for systematic recording of dietary data during interview (Tapsell LC et al, 2000). The consumption of food items were recorded in chronological order from morning till dinner time. For Nutrient intake calculation national institute of nutrition book: nutritive value of Indian foods (Gopalan BV, et.al, 2007) and laboratory estimates of nutrient in cooked food items (Chiplonkar SA and Agte VV, 2007) was used.

Starting with the first meal of the day, participants were asked to describe their usual eating patterns during pregnancy over the last three months and for GDM mothers before diagnosis and to qualify the amount with details of how much and how often each foods was consumed. Cost-effectiveness and low respondent burden, the 24 hour diet recall was the preferred method for dietary assessment of pregnant women, however the accuracy of a 24 hour diet recall is dependent on the respondent recollection of the types and quantities of food, beverages and supplements they were consuming, hence the dietary protocol was improved to increase reporting accuracy by using portion sizes.

Dietary habit questionnaire was used, to calculate amount of non-vegetarian food consumption per day and amount of PUFA, MUFA, ω3, ω6 from visible fat (oil) intake per day was (Appendix).
Plasma vitamin C levels were assessed to validate the results derived from SFFQ and 24 hour diet recall. In addition so-called validity was assessed among women by comparing nutrient results from the SFFQ with 24 hour diet recall collected at random and non-consecutive days (Saldana TM, et.al, 2004).

Also section ‘I’ of questionnaire (Dietary habits) was used to get information for this part:

This part was starting with subjects’ type of diet, which included of complete vegetarian, lacto-vegetarian (only taking milk), ovo-vegetarian (only taking egg), lact-ovo vegetarian (taking both egg and milk), non-vegetarian (twice per week, intake of any of non-vegetarian sources), (dietetics, Srilaxmi, 2007) and if any cases was non-vegetarian during occasions. Then if the subject was non-vegetarian the frequency of non-vegetarian sources consumption with the portion size was asked. It means the mother were asked how many times per day, week, month consuming non-vegetarian sources like chicken, fish, mutton, beef, etc. then on the basis of daily frequency, if the subject consumed twice per week, the frequency would be 2/7 which means 0.28.

Since at the end of first trimester it is very usual to take nutritional supplements, these questions also added about iron, folic acid, calcium, any other vitamins, by the generic/commercial name of the drug, number of tablets and daily frequency of taking. Also in this part of questionnaire time of starting supplements were covered. If mother was taking any other nutritional supplements like born vita, proteinex, complan, or any Ayurvedic medicine like shatawari, that also was recorded, since all these supplements are extra sources for getting vitamin C.

Dietary habits questionnaire was used to assess the consumption of foods like type and amount of oil, sugar, Jaggary. The frequency of consumption was calculated. Per capita consumption was used wherever needed. Since most source of LCPUFA (ω3) is from marine sources and most of Indian population is vegetarian, hence in present study we assessed EFA levels in oil consumption.

Dialing into the digital age, food transition habits are adopting more among individuals and can consequent different health problems in long term. Hence a part of dietary questionnaire specified to calculate score of frequency for consumption of these
food groups during pregnancy as well as frequency of eating outside according to the method explained above. These food groups like fast foods (pizza, burger, vadapav, etc), ready to eat foods (magi noodles, magi soups, etc), sweet snacks (chocolate, sweetened fruit juices, cola beverages, jellies, etc) fried salty snacks (cutlet, burger, chips, etc); bakery products (butter cake, dough nuts, cream roll, etc).

To estimate nutrient intake there is a need for a food composition table listing the average nutrient content of food staffs in SFFQ and 24 hours diet recall. The food composition data base was used to convert intake of food into nutrients namely carbohydrate, protein, fat, energy, percent of carbohydrate, protein, fat from calorie, vitamin C, iron, fiber. For this purpose Nutrient values of Indian food (Gopalan. C 2004), USDA data base for HORAC foods, Chiplonkar et.al 2007 and Agharkar food data base were used.

For supplements and readymade foods it was calculated by food label information and the proportion and frequency mother was taking (e.g. one Marie gold biscuit equal to 6gm (Agarkar1999)). The “adequacy intake percentage of each subject for each “were calculated using recent Indian RDA data base, 2010.

Two diet recalls, comprising of one working day and one weekend used to obtain overall average consumption. The standard portion size katories, spoons, rotis (Indian bread) (small, medium, large) were used for each food staff. To calculate 24 hour diet recall if the mother was taking any supplements according to company and portion size (number of tablets, scopes, spoons, etc) micronutrient level were calculated and considered in total daily intake.

Nutrient density was calculated by nutrient intake*100/RDA. The mean intake of different food stuffs consumed was computed for a day and compared with the balanced diet.

Nutrient adequacy ratio (NAR) were calculated from the following formula (Anita Malhotra, 2007): Subject’s nutrient intake per day/RDA of that nutrient

Adequacy percentage was calculated by multiplying the above formula to 100 (Goerge J, 2003).
In order to assess the diet quality, the adequacy of nutrient intake by each subject was computed in terms of Nutrient Adequacy Ratio (NAR) using \( \text{NAR} = \frac{\text{Subject's nutrient intake of a day}}{\text{RDA of the respective nutrient}} \). Thereafter, the subjects were categorized as those having an adequate (\( \geq 1.00 \)), fairly adequate (\( 0.66 < \text{NAR} < 1.00 \)) or in-adequate (\( \text{NAR} < 0.66 \)) NAR for various nutrients. (Anita Malhotra, et al., 2007) Since NAR is not a good indicator for assessing the adequacy or inadequacy of energy intake therefore, energy intake data were expressed as percent of RDA for the particular age group.

3.5. Biochemical measures and Plasma vitamin C status:

The method is adapted from Joseph H Roe, Carl A Kuether 1942, with some modifications (George J, 2003).

![Biochemical measures and Plasma vitamin C status](image)

Figure 3.2. Plasma vitamin C analysis steps

Figure3.2. Plasma vitamin C analysis steps
Methods:

1. **DNPH reagent** (2, 4-Dinitrophenylhydrazine): 2gm 2,4DNPH was mixed with 250 mg thiourea and 30 mg CuSO4, 5H2O and 100 ml of 9N H2SO4 and stored in refrigerator.

2. **65% H2SO4**: 65 ml concentrated H2SO4 to 35 ml distilled water were solved.

3. To quantitate vitamin C concentration, the frozen samples were thawed at room temperature.

4. **Blank tubes**: 1 ml of 10% TCA was added to 0.4 ml DNPH reagent. To prepare standard test tubes, 0.8 ml of 10% TCA with 0.4 ml DNPH reagent and 0.2 ml standard ascorbic acid (2mg/dl) were mixed together. For each sample, 1 ml of thawed supernatant was added to 0.4 ml DNPH reagent. Samples were incubated at 37° C for three hours. Samples then chilled on ice. After this 1.6ml of cold 65 % H2SO4 was added to each sample. Samples were incubated at room temperature for 30 minutes. Samples were read at 520 nm on spectrophotometer. To calculate the amount of plasma ascorbic acid from absorbance on spectrophotometer, the following formulas were used: Ascorbic acid (mg/dl) = (OD of test/OD of standard) × concentration of standard ascorbic acid, and then to convert mg/dl to μmol/l: mg/dl × 56.77=μmol/l.

5. **Standard ascorbic acid** (2mg/dl): 20 mg vitamin C was mixed with 100 ml distilled water and was further diluted 10 times to make working standard. This solution was prepared freshly every time.

Protocol:

First two ml non-fasting blood was drawn from the patients’ vein in EDTA tubes (at the time of 1hour PP of OGTT test). As soon as the blood collected, it was being transferred to lab on ice and away from UV.

The samples were spun at 3600 RPM for 12 minutes at 4 °C. And plasma was separated. 0.4 ml plasma was mixed with 1.6 of ml 10% TCA (To make 10% TCA, 1 ml of 100% TCA was mixed with 9 ml distilled water.) sample was spun again at 8000 RPM for 10 minutes. Pellet was discarded and clear supernatant was stored at -80 till further analysis. To ensure accuracy duplicate measurements was performed for all samples.
Table 3.2: measurement of plasma vitamin C at a glance

<table>
<thead>
<tr>
<th></th>
<th>Blank</th>
<th>Standard</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test supernatant</td>
<td>-</td>
<td>-</td>
<td>1ml</td>
</tr>
<tr>
<td>Standard ascorbic acid(2mg/dl)</td>
<td>-</td>
<td>0.2ml</td>
<td>-</td>
</tr>
<tr>
<td>10%TCA</td>
<td>1ml</td>
<td>0.8ml</td>
<td>-</td>
</tr>
<tr>
<td>DNPH reagent</td>
<td>0.4ml</td>
<td>0.4ml</td>
<td>0.4ml</td>
</tr>
</tbody>
</table>

Incubate at 37 °C for 3 hours Then chilled in ice bath

<table>
<thead>
<tr>
<th>65%H2SO4 cold</th>
<th>1.6ml</th>
<th>1.6ml</th>
<th>1.6ml</th>
</tr>
</thead>
</table>

Incubate at room temperature for 30 minutes

Optical density 520 nm

Section D of questionnaire: Biochemical parameters:

Women were tested for GDM between 24-28 weeks of gestation according to international guidelines (Metzger BE, et.al, 1991)

Since the pregnant subject had regular hematological check up, her hemoglobin and RBC count were recorded using laboratory reports from second trimester.

To measure Plasma vitamin C blood sample from the ante-cubital vein was collected in an EDTA tube vacutainer and method of Joseph H Roe and Carl A Kuether 1942, was used with little modification. (The method is explained elaborately in the following sections). All methods of plasma vitamin C determination were performed without knowledge of pregnancy outcome.

OGTT test was done and the results recorded. For this purpose the subject was asked to be on fast from the night before test. Fast was meant no eat, drink or smoke after midnight. However the subject was allowed to drink one glass of water in the morning. After fasting blood collection, subjects were asked to drink 75 gm glucose drink (75gm
unhydrous glucose with 300 ml water). A new blood test then was taken one and two hours later. If mother was diagnosed as diabetic she was referred for further consultation.

Screening for GDM is important because studies have shown that treating even mild GDM reduces morbidity of both the mother and newborn (Landon et al, 2009). Universal screening for GDM detects even mild cases and improves maternal and offspring prognosis.

**Glycemic status** evaluated and GDM was confirmed by a 75 g oral glucose tolerance test (OGTT) using IADPSG (International association of diabetes and pregnancy study group) criteria, using a threshold of fasting plasma glucose (FPG) 92mg%, 1hour 180mg%, 2hour 153mg%. The method was explained elaborately in above.

### 3.6. Physical activity:

Physical activity score were calculated using a standardized and international questionnaire (appendix) (IPAQ, short format, www.ipaq.ki.se, completed May 2001).

Information regarding usual physical activity of mother prior to present study was obtained by using interview technique. The mother was asked the type of activity and duration of each activity during at least seven days prior to confirmation of present pregnancy.

Data from the short IPAQ questionnaires were summarized according to the physical activities recorded (walking, moderate, and vigorous activities) and estimated time spent sitting per week. Note that the sitting questions were developed as separate indicators and not as part of the summed physical activity score (Craig CL, et.al, 2003). The weighted MET-minutes per week (METmin-wk⁻¹) were calculated as duration × frequency per week × MET intensity, which were summed across activity domains to produce a weighted estimate of total physical activity from all reported activities per week (MET·min·wk⁻¹). For brevity and clarity of presentation, only the total physical activity MET-minutes per week and total minutes per week in sitting activities are reported here (Craig, CL et.al, 2003). (First MET values were derived from indirect caloriemetry (Craig CL, et.al, 2003)). The formulas were as follows:

\[
\text{Vigorous (minute/week)} = 8 \times \text{vigorou activity (minute)} \times \text{vigorou activity days}
\]
Moderate (minute/week) = 4 × moderate intensity activity (minute) × moderate intensity activity days
Walking (minutes/week) = 3.3 × walking (minute) × walking days
Total physical activity MET (minute/week) = vigorous + moderate + walking

Sitting calculated and presented as minutes per one of week days.

As there are no established threshold for presenting MET-minute the IPAQ committee decided that data should be reported as comparisons of median values and inter-quartiles ranges for different population, however Craig CL et.al in a study to validate IPAQ validate in 12 countries, concluded that Considering the diverse samples in this study, IPAQ has reasonable measurement properties for monitoring population levels of physical activity among 18- to 65-year-old adults in diverse settings.

Only values of 10 or more minutes activities was included in the calculation. Scientific evidence indicates that episodes of at least 10 minutes are required to achieve health benefits, responses less than 10 minutes was considered as zero (Chasan-Taber, et.al, 2005).

3.7. Stress:

In this study, stress was known as confounder variable. To evaluate stress level, an international standardized questionnaire (SRRS, social readjustment rating scale) was used (Holmes TH & Rahe RH, 1967). This scale was retested by Rahe, et.al in 1970 and also in the year 2000 as a predictor of illness. In addition this scale was assessed cross-culturally, in Japanese and Malaysia (Masuda M et.al, 1967) and Malaysian Americans (Woon TH et.al, 1971) consequently Considering the diverse samples in these studies, Rahe and Holmes life stress questionnaire has reasonable measurement properties for monitoring population levels of stress among 18- to 65-yr-old adults in diverse settings, even though many researchers have expressed concerns about its usefulness.

However questionnaire was modified to suit the present study population. There are two types of events, positive like increasing salary, marriage, etc and negative events like divorce, loss of job, etc. Also one event has different effect on different people, e.g. leaving a job would be disaster for a person whereas would be a better opportunity for the other. Here the underlying assumption was that the negative nature of events is not the
important factor, but the amount of change that is required to readjust to a tolerable level of functioning. Any change, whether desirable or not, was seen as stressful.

To measure stress according to Holmes and Rahe stress scale, the number of life change units that applied to events in the past year of an individual’s life are added and the final score was given a rough estimate of how stress affect health. Recently this questionnaire was used to assess stress in Iranian population to estimate maternal stress in pregnancy and umbilical cord Ig. E (Bidaki R, et.al, 2011).

The subjects were asked that “during the last twelve months has she had any of the following things happen to her? If so, simply circle the number following those items which has occurred in her life recently (and only those items that apply to her)”.

Stress category was determined using following categories:
1. the score greater than 300, major life stress
2. 200-299 moderate life stress
3. 150-201 mild life stress
4. <150 low life stress

Since this part comprised of some delicate questions, coded questionnaire was provided separately to the subjects, to keep confidentiality.

As there are no established thresholds for presenting stress score data was reported as comparisons of median values for this population.

3.8. Passive smoking:

There were three questions in the questionnaire for this purpose. The questions including whether the mother is exposed to any person with smoking habits at least once per day, if the response was yes, they should specify how many times per day.

3.9. Outcomes:

In this section data regarding any pregnancy related problem, like fetal macrosomia, poly Hydraminos, IUGR (Intra Uterine Growth Retardation) was recorded.

Ultrasonography is important for monitoring diabetic pregnancies and potentially improving both prenatal management and fetal outcome. It is used to assess four major factors: Gestational age, Congenital anomalies, Growth abnormalities, Macrosomia, Intrauterine growth restriction, fetal well-being (dynamic assessment).
In this study all ultra-sonographies (USG) reports available with the subject till the time of interview were recorded which provided date and weeks in which that USG was done. The reports provided the data regarding fetus head circumference (HC), femur length (FL), biparital diameter (BPD), abdominal circumference (AC), abnormality if any in the placenta and amniotic fluid volume (AFV) and the velocity of fetal growth according to gestational age, etc. Polyhydraminos was defined as Amniotic fluid index (AFI) greater than 25cm (or above 95th percentile) or more than 2000 ml amniotic fluid (American pregnancy association). (Section E of questionnaire: Antenatal ultra-sound reports).

Birth weight of the baby was obtained from medical record of hospital. Also SGA (small for gestational age), LGA (large for gestational age), AGA (appropriate for gestational age) category in which the neonate will included, were recorded from the same. To calculate gestational age of neonates Ballard maturational assessment was used (Section F of questionnaire: Maternal present obstetric information and any complication, appendix).

SGA is defined as weight below 10th percentile (Erika Aaron, et.al, 2012), whereas weight above 90th percentile and weight between 10 and 90th percentiles demonstrates LGA and AGA respectively (Nelson textbook of pediatrics, 2011). LBW or low birth weight is referred to neonate lower than 2500 g at birth (Rao, et.al, 2001).

Macrosomia is a condition of fetal over-growth leading to a LGA fetus. It is defined as birth weight > 4000 grams or above 90 the percentile for population and sex specific growth curves. It’s commonly seen in GDM, prolonged pregnancies and pregnancies complicated with pre-existing diabetes mellitus (Kale SD, et.al, 2005, Europe pub med central.)

IUGR is defined as birth weight below 10th percentile of birth weight for gestational age reference curve, as determined through an ultrasound. This can also be called SGA or fetal growth restriction (De-Onis M. et.al, 1998).

Pre-term delivery defined as delivery sooner than 37 weeks of gestation (Kale SD, et.al 2005).
In Section G of questionnaire (Maternal delivery and post partum information) following data was recorded from delivery reports of the mother and medical birth registry available in the medical records of the hospital.

1. Gestation weeks and days at the time of delivery,
2. Mode of delivery. In case of caesarean whether elective or emergency.
3. The outcome, live birth, still birth or neonatal death.
4. Sex of baby

Neonatal complications like hypoglycemia, hyperbilirubinemia, Respiratory Distress Syndrome (RDS), neonatal mortality, admission to NICU (if any), were recorded. Section H of questionnaire, was included of neonatal complications.

**Neonatal hypoglycemia** has been defined as plasma glucose level less than 44 gm/dl in the first few days after birth (Seshiah V, 2012).

**Neonatal hyperbilirubinemia** which occurs mostly in term babies will be diagnosed as Icterus, which however, becomes apparent on the skin when serum bilirubin reaches more than 5 mg/dl. “Almost all neonates (60% Term and 80% Preterm) will have bilirubin greater than 5 mg/dl in the first week of life and about 6% of term babies will have levels exceeding 15 mg/dl.

**Respiratory distress syndrome** (RDS) or HMD hyaline membrane disease occurs mostly in premature deliveries (Alday L, et al, 2006). Lillian defines RDS as follows:

“When an infant is born prematurely, they lack adequate amount of surfactant production to form in the alveoli causing the lungs to collapse and making it very difficult for the baby to get enough air. Although surfactant production starts near 22 weeks gestation, the pulmonary system is easily disrupted by hypoxemia, hypothermia, and acidosis, all of which plaque the premature neonate. It is not until the mature surfactant is produced near 35 weeks gestation that the above stressors do not disrupt the production and the fetal lungs are considered mature.”

To address research questions in present study statistical analysis comparison, descriptive and inference statistic was performed. The significance threshold of $P=0.05$ was used in all analyses. All statistical analyses were performed using statistical package for social sciences (SPSS IBM version 21) software.

At the commencement, double data entry with overlap for data verification was done and for this purpose data were recorded in a standard questionnaire. Overall information (i.e. explanatory and outcome) including exposure data were entered into two different excel sheet (double data entry) by two different person. This allowed us to check for any variation in the data sheet using the following Excel function 

```
=IF (EXACT (SHEET1!A2, SHEET2!A2), 0, SHEET1!A2&"/"&SHEET2!A2)
```

Checking for outliers (out of range values) and missing data was accomplished. Less than 12% of plasma vitamin C values were not available (24 out of 200), but since it was from control group, number of subjects included in the study were higher than required; hence it did not affect the results.

For skewed data log transformation was undertaken for conversion to normality.

**Preliminary data analysis** including Descriptive statistic, comparison of means by independent t-test, Mann-witney test was done and examining the association between variable of “predictor of interest (POI) and outcome variable also verified.

**Descriptive analysis** was accomplished to express means, standard error of mean, standard deviation and minimum and maximum of continues variables that consisted of maternal macronutrient and micronutrient intakes, percentage of energy derived from macronutrient intake, expectant mothers age, maternal pre-pregnancy BMI and physical activity score, maternal stress score, maternal hemoglobin, RBC counts, plasma vitamin C and glucose level and Neonatal birth weight. Subsequently we cross tabled the explanatory variables, by the outcome variables, using the Chi-square test ($\chi^2$) for categorical variables and the independent t-test when data was continuous. In other words the characteristics of pregnant GDM were compared to those non-GDM using t-test for continuous variable and $\chi^2$ test for categorical data. Also the distribution of continues variables were categorized and re-compared.

114
Cumulative frequency percentage of nutrient also calculated and showed in graphical format.

**Uni-variate analysis** of all variables and examining correlations (Pearson, spearman) between exposure and outcome variables was another method used for statistical analysis in this study.

Modeling mean by **linear regression** to find out the predictor variables for maternal plasma glucose levels, premature delivery and LBW.

Also Women were grouped according to tertiles determined by distribution of maternal plasma vitamin C among the entire cohort. Distribution of clinical, biochemical and nutritional characteristics of participants were also examined according to these quartiles.

Confounders were assessed by **partial correlation**, with and without adjustment. That means to assess confounding; variables were entered into a generalized linear model, one at a time. Covariates specifically evaluated as confounders included age, stress, and household income.

Graphical presentations, report tables, interpret the results and helped for data visualization.

**Data safety and confidentiality**: We always had up to date anti-virus software installed on machine and the data sets regularly backed up. Each study subject provided a unique identity code for confidentiality and the final clean the database was then transformed and stored in SPSS software version 21.0 for further analysis.
Table 3.3. Selection of statistical tests for description of variables and hypotheses testing, which one and why?

<table>
<thead>
<tr>
<th>Target</th>
<th>Quantitative/Normal</th>
<th>Qualitative/Non-normal</th>
<th>Bivariate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>Mean(SD)</td>
<td>Medain (25th, 75th percentile)</td>
<td>-</td>
</tr>
<tr>
<td>Compare two independent group</td>
<td>Independent t-test</td>
<td>Mann-whitney</td>
<td>Chi-square</td>
</tr>
<tr>
<td>Correlation of two variables</td>
<td>Pearson correlation coefficient</td>
<td>Spearman correlation coefficient</td>
<td>Bivariate correlation coefficient</td>
</tr>
<tr>
<td>Prediction of one variable according to two or more variables</td>
<td>Linear regression simple/multiple</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Test of normality</td>
<td>Kolmogroph Smirnoph</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusting for confounders</td>
<td>Partial correlation</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In brief the data analyses were performed into two parts. The first part of analysis was conducted to uncover the relationship between exposure and outcome variables. This analysis was coincided to the type of data means, chi-square test ($\chi^2$) for qualitative variables. Independent t-test and multiple linear regressions model for quantitative.

Categorical variables were consist of occupation, education level, monthly income, family type, history of diabetes in family, history of bad obstetric history, taking any kind of supplements, neonatal complications, type of diet, smoking habits.

Quantitative variables including, maternal pre-pregnancy height, weight physical activity score, life stress score, maternal micro and macro nutrient intake, gestational weeks at delivery, plasma level of vitamin C and glucose, RBC counts and hemoglobin level.

**Data management:**

The normal distribution of variables was checked by Kolmogroph Smirnoph test, histogram. Log10 transformation and square root for non-normal variables was performed. Again checking for normality was carried out. The following figures show the changes.
Figure 3.3.: Histogram of dietary vitamin C, before (left) and after (right) transformation (above and following respectively):

Figure 3.4: Histogram of one hour PG, before (left) and after (right) logarithm transformation