Students were asked to analyze the data on "**Identifying Problem Pregnancies in the Developing World"**. The questions of interest were stated in the documentation as

1. Is there evidence that weight profiles and/or SFH profiles over pregnancy differ between women who do and do not deliver SGA babies?

2. Is it possible, using measurements taken prior to week 30 of pregnancy, to develop a model which accurately distinguishes between women who will and will not have growth retarded babies?

In assigning the project, I asked students to concentrate primarily on the first question, and then I stated: "You do not need to create a predictive model hinted at in the second listed question *per se*, but you should comment on the likelihood that the associations you found in answering the first listed question could be used to provide sufficiently accurate predictions."

The statement of the problem contained the following information that I regard as highly relevant to my analysis of the data:

- Low birth weight of babies can be due to pre-term delivery, intra-uterine growth restriction (IUGR), or both.
- Measures of SFH tend to increase approximately linearly with gestational age.
- In non-obese women, the SFH in centimeters tends to very close to the GA. In obese women, this correspondence is not as good, though there still tends to be linear increase with GA.
- Smoking is associated with IUGR.
- Age was not explicitly mentioned as a risk factor for IUGR, though hypertension and other circulatory problems were. These, of course, increase in prevalence with age.
- Parity was not explicitly mentioned as a predictor for IUGR, though the possibility of maternal factors that would be common across pregnancies was mentioned.

Additionally, in the session devoted to discussion of the project, I discussed the role that height might play as

- an indicator of general nutritional status, and / or
- an indicator of ethnicity: the Khoi-San (Bushmen) in South Africa tend to be of short stature, and it might seem reasonable that their babies would be similarly smaller.
    - I note that I am uncertain the degree to which that shorter stature might be genetic or might be indicative of long standing nutritional status among those peoples. Certainly height distributions in other parts of the world have changed dramatically in response to what I presume to be environmental factors.

The data was provided in two files:

- One file contained information about the mother (age, height, parity, smoking) and the outcome of the pregnancy (baby's sex, birthweight, gestational age at birth, and an indicator that the baby was small for gestational age (SGA).
- One file contained longitudinal data from the prenatal clinic visits, including the estimated gestational age, mother's weight, and symphysis to fundal height (SFH).

Pertinent information (i.e., information that would be germane) that was not provided in the documentation includes:

- The sampling scheme was not adequately described. In particular, it is not clear exactly what were the eligibility criteria for the study beyond singleton pregnancy being followed at some unspecified antenatal clinic. The planned schedule of antenatal visits was completely unspecified, and there was no discussion of the possibility that more frequent monitoring would have been prescribed for women whose weight gain, SBP, or SFH measurements were aberrant.
- It is not explicitly stated that no enrolled woman was ever transferred to more intense monitoring at another clinic. Were that to have happened, our sample is probably fatally flawed.
- How the estimated GA was determined in the antenatal clinics was not described. Presumably some patients do have an ultrasound to estimate the GA at some point during their pregnancy. Information from the internet suggests that in South Africa they do try to schedule a single ultrasound to establish more reliable estimates of GA at some point during pregnancy (18-20 weeks would be a good time to do so), but ultrasound is not universally available (some clinics use itinerant ultrasonographers, others have no such service). If I was asked, I told the students they could presume it was LMP.
- How the GA was estimated at birth was not described. As noted below, these often conflict with the measurements made during antenatal visits (the GA estimated at birth is sometimes less than the GA estimated at some antenatal visits). If asked by students I told them that they could regard that these were errors or differences in measurement. In fact, it is standard practice to estimate GA at birth using criteria that are different from what might have been used in prenatal visits. Generally, in the US it would be regarded that ultrasound in the 18-20 week period would provide the best estimate of GA, and the measurements made after birth would be less reliable. However, if prenatal GA is estimated only by LMP and SFH, the measurements made at birth might be regarded as more precise (I was taught that the leading cause for too large for GA is wrong dates of LMP. For instance, there is often some slight bleeding at the time of implantation, and that can be mistaken for a light period. In women with irregular menstrual cycles, the LMP would be obviously less reliable.)

Below I present results and observations from a very cursory analysis of these data. I use SGA as an example, imagining that similar analyses could be done with LBW and pre-term delivery. Of course, ultimately a similar action (referral to a high risk clinic) is to be taken for subjects at high risk for any of these adverse outcomes. Hence, a strong argument could be made for defining a variable indicative of women falling into at least one of the adverse pregnancy outcomes. Analysis of that endpoint would proceed in the same fashion.

**Description of the Sampling Scheme**

Data is provided on 755 mothers, though there is some missing data:

- height is missing for 6 mothers
- smoking status is missing for 4 mothers
- birthweight of the baby is missing for 4 pregnancies
- sex of the baby is missing for 4 pregnancies
- gestational age at birth is missing for 5 pregnancies

Note that we do not have information on the pre-pregnancy weight of women, but I am interested in having some measure of obesity. I thus computed BMI using the first clinic visit for each woman. As noted below, there is some variability in the GA at time of enrolment, but I did not try to make any adjustment for that when computing the BMI for the use in this analysis.

From a linear regression analysis of the GA at enrolment as a function of age, bmi, smoking status, parity, and interactions between age and smoking and between age and parity, there was at most a slight trend for women with higher BMI to have been enrolled later. Some of this could be an artifact of the use of the weight during pregnancy, but as the interquartile range for GA at enrolment was 20-24 weeks, this was probably not too much of an issue.

The following table provides the distribution (mean, SD, min – max, sample size) of estimated GA at time of enrolment as stratified by women whose babies were ultimately declared SGA. The GA at enrolment was fairly comparable between SGA and non-SGA.

| | sga.0 | sga.1 | sga.ALL |
|---|---|---|---|
| Enrol GA | 22.6 (4.14; 18.0-39.0, n=650) | 21.9 (3.36; 15.0-36.0, n=105) | 22.5 (4.04; 15.0-39.0, n=755) |
| Number Visits | 7.85 (2.20; 2.00-14.0, n=650) | 7.11 (2.68; 2.00-13.0, n=105) | 7.74 (2.28; 2.00-14.0, n=755) |

Data is provided on 5,849 prenatal visits.

- The estimated GA at the visit is missing for two such visits, though for each of these mothers there is another visit with the same weight and SFH measurements that do have the GA recorded.
- There are 9 prenatal visits missing weight data.
- There are 13 prenatal visits missing SFH data.

The above table also provides the distribution (mean, SD, min – max, sample size) of the number of prenatal visits for each enrolled woman as stratified by SGA group. Again, the number of visits is roughly comparable.

From a linear regression analysis of the number of prenatal visits as a function of age, bmi, smoking status, parity, gestational age at birth, and SGA status and interactions between age and smoking, between age and parity, and between SGA and gestational age at birth, we find that, as might be expected, women who delivered their baby after longer gestations tended to have more antenatal visits. Multiparous women tended to have slightly fewer visits.

Importantly, there did not seem to be a marked difference between the number of visits for women who did or did not have a baby judged to be SGA. (Of course, quality of care issues would dictate that those women should be seen more often. However, statistical issues that might arise due to biased sampling do not appear to predominate in this data.) In more detailed analyses considering the odds of a woman being seen at the clinic

for each estimated GA between 18 and 40 weeks (adjusted for age, bmi, smoking, parity, gestational age), there did not appear to be any trend by the babies' ultimate SGA status. (Data from these 23 logistic regressions not shown.)

**Unusual Measurements in the Data**

I examined the longitudinal weight data on the women to look for possible data entry errors (and/or signs of pre-eclampsia). The criterion I used was to identify those women whose weight change between successive visits would be at a rate of 2.5kg or more per week. At least 33 women in the data have successive weight measurements suggestive of errors in the data (or a very severe amount of fluid retention). As some of these women showed differences corresponding to 10kg per week, I suspect that there is a high amount of error in data collection and/or data entry. In the presence of such high error rates, our ability to detect trends in weight gain between SGA and non-SGA pregnancies is greatly diminished.

Furthermore, as discussed in class, I serendipitously found records for two distinct maternal ID codes that I believe to be the same person.

**Descriptive Statistics for Enrolled Mothers**

The following table presents statistics describing the distribution (mean, SD, min – max, n) for the women enrolled in the study stratified according to whether their baby was deemed to be SGA. As might be expected, the prevalence of smoking is higher among those women whose babies were deemed to be SGA.

| | sga.0 | sga.1 | sga.ALL |
|---|---|---|---|
| age | 24.9 (5.45; 14.0-43.0, n=650) | 23.8 (4.90; 16.0-35.0, n=105) | 24.8 (5.39; 14.0-43.0, n=755) |
| height | 157 (6.54; 106-176, n=650) | 155 (5.87; 142-172, n= 99) | 157 (6.50; 106-176, n=749) |
| bmi | 25.7 (4.59; 16.9-49.1, n=650) | 24.0 (4.29; 15.9-39.6, n= 99) | 25.4 (4.58; 15.9-49.1, n=749) |
| parity | 1.13 (1.23; 0.0-6.00, n=650) | 0.90 (1.11; 0.0-6.00, n=105) | 1.10 (1.21; 0.0-6.00, n=755) |
| smoker | 28.7% (n=647) | 43.3% (n=104) | 30.8% (n=751) |
| male | 52.4% (n=647) | 42.3% (n=104) | 51.0% (n=751) |
| enrol | 22.6 (4.14; 18.0-39.0, n=650) | 21.9 (3.36; 15.0-36.0, n=105) | 22.5 (4.04; 15.0-39.0, n=755) |

**Descriptive Statistics for Pregnancy Outcomes**

The following table presents statistics describing the distribution (mean, SD, min – max, n) for the pregnancy outcomes for women enrolled in the study stratified according to whether their baby was deemed to be SGA. SGA comprised 13.9% of the population, and the SGA babies had lower average birthweight, with 72.1% of them meeting criteria as low birth weight (LBW = < 2500 g). Interestingly, none of the non-SGA babies were deemed to be LBW. This suggests to me that improper criteria were used to define SGA, as we actually would expect only about one-third of LBW to be SGA (a preterm infant (GA < 38 weeks) can still be appropriately sized for GA, as there are other reasons that a woman might deliver early).

| | sga.0 | sga.1 | sga.ALL |
|---|---|---|---|
| sga | 0.0% (n=650.0) | 100.0% (n=105.0) | 13.9% (n=755.0) |
| bweight | 3246 (402; 2510-4730, n=647) | 2231 (412; 1035-3780, n=104) | 3106 (534; 1035-4730, n=751) |
| lbw | 0.0% (n=647) | 72.1% (n=104) | 10.0% (n=751) |
| gesage | 39.4 (1.24; 38.0-44.0, n=647) | 37.9 (2.20; 30.0-42.0, n=103) | 39.2 (1.50; 30.0-44.0, n=750) |
| preterm | 0.0% (n=647) | 37.9% (n=103) | 5.2% (n=750) |

**SFH Profiles by SGA Status**

The following table presents statistics describing the distribution (mean, SD, min – max, n) for the SFH measurents made at visits corresponding to GA of 18 – 40 weeks for women enrolled in the study. In each case, the statistics are stratified according to whether their baby was deemed to be SGA. Apparent from this table is a tendency for SFH measurements to be similar up to about 25 weeks GA, and then there is increasing separation between the groups as the pregnancies progress. It should be noted, that the SGA group does tend to deliver earlier than the non-SGA group, so sample sizes in the SGA group will be diminished for that reason.

| | sga.0 | sga.1 | sga.ALL |
|---|---|---|---|
| GA18 | 16.9 (1.60; 13.8-22.3, n=86.0) | 16.7 (1.73; 14.0-19.0, n= 9.0) | 16.9 (1.61; 13.8-22.3, n=95.0) |
| GA19 | 18.2 (1.91; 12.0-23.7, n=77.0) | 19.0 (4.27; 14.0-28.0, n=10.0) | 18.3 (2.28; 12.0-28.0, n=87.0) |
| GA20 | 19.1 (1.79; 12.0-28.9, n=144) | 19.2 (0.94; 18.0-21.0, n= 15) | 19.1 (1.73; 12.0-28.9, n=159) |
| GA21 | 20.2 (1.55; 15.5-24.1, n=76.0) | 19.6 (1.79; 17.0-25.0, n=16.0) | 20.1 (1.60; 15.5-25.0, n=92.0) |
| GA22 | 21.2 (1.75; 16.8-26.6, n=154) | 21.3 (1.46; 19.0-25.0, n= 26) | 21.2 (1.71; 16.8-26.6, n=180) |
| GA23 | 22.1 (1.68; 17.0-28.4, n=104) | 22.5 (2.23; 19.0-28.0, n= 21) | 22.1 (1.78; 17.0-28.4, n=125) |
| GA24 | 23.3 (1.82; 18.1-30.0, n=165) | 23.4 (2.84; 19.0-34.0, n= 36) | 23.3 (2.03; 18.1-34.0, n=201) |
| GA25 | 24.3 (1.64; 17.6-29.8, n=126) | 23.7 (1.77; 20.0-28.0, n= 25) | 24.2 (1.67; 17.6-29.8, n=151) |
| GA26 | 25.3 (1.63; 20.5-30.5, n=194) | 25.2 (1.69; 21.0-29.0, n= 25) | 25.3 (1.63; 20.5-30.5, n=219) |
| GA27 | 26.3 (1.70; 20.9-32.5, n=123) | 25.8 (2.20; 22.0-30.0, n= 26) | 26.2 (1.80; 20.9-32.5, n=149) |
| GA28 | 27.4 (1.65; 21.2-32.3, n=307) | 26.5 (2.46; 22.0-35.0, n= 50) | 27.3 (1.81; 21.2-35.0, n=357) |
| GA29 | 28.4 (1.68; 23.8-33.4, n=123) | 27.7 (2.65; 22.0-35.0, n= 26) | 28.3 (1.89; 22.0-35.0, n=149) |
| GA30 | 29.3 (1.67; 23.7-36.5, n=349) | 28.3 (1.80; 25.0-34.0, n= 50) | 29.2 (1.71; 23.7-36.5, n=399) |
| GA31 | 30.1 (1.87; 25.8-37.4, n=140) | 29.4 (2.20; 26.0-37.0, n= 29) | 30.0 (1.94; 25.8-37.4, n=169) |
| GA32 | 31.1 (1.63; 26.0-35.6, n=371) | 30.0 (2.29; 25.0-36.0, n= 51) | 31.0 (1.76; 25.0-36.0, n=422) |
| GA33 | 31.8 (1.83; 26.0-37.8, n=187) | 30.4 (2.74; 22.0-35.0, n= 30) | 31.6 (2.03; 22.0-37.8, n=217) |
| GA34 | 33.0 (1.79; 26.8-39.4, n=376) | 31.7 (2.42; 25.0-40.0, n= 59) | 32.8 (1.93; 25.0-40.0, n=435) |
| GA35 | 33.8 (1.77; 29.3-40.0, n=200) | 31.8 (2.24; 27.0-36.0, n= 32) | 33.5 (1.96; 27.0-40.0, n=232) |

| | | | |
|---|---|---|---|
| GA36 | 34.6 (1.72; 29.9-42.4, n=422) | 33.0 (2.46; 28.0-44.0, n= 58) | 34.4 (1.90; 28.0-44.0, n=480) |
| GA37 | 35.4 (1.84; 30.6-41.5, n=431) | 33.7 (1.88; 30.0-37.0, n= 53) | 35.2 (1.92; 30.0-41.5, n=484) |
| GA38 | 35.9 (1.97; 31.3-43.5, n=416) | 34.0 (1.93; 30.0-37.0, n= 39) | 35.7 (2.04; 30.0-43.5, n=455) |
| GA39 | 36.3 (2.25; 30.6-45.5, n=297) | 33.9 (2.17; 28.0-38.0, n= 24) | 36.1 (2.33; 28.0-45.5, n=321) |
| GA40 | 36.7 (2.21; 31.6-43.0, n=161) | 34.2 (2.17; 31.0-39.0, n= 13) | 36.6 (2.30; 31.0-43.0, n=174) |

**Anticipating Models Prognostic for SGA Based on Data at 30 Weeks GA or Before**

In considering the models that might be used in a low technology environment to screen for pregnancies at high risk for SGA, I noted the following:

- Causes of IUGR are primarily thought to be circulatory, nutritional, infectious, or genetic. *A priori* we would expect any of these to manifest themselves on SFH and weight gain later in pregnancy as the fetus gains weight.
- Any prognostic model would likely be used sequentially. That is, when a woman's prenatal visit was suggestive of IUGR, that woman would be referred to another clinic. Hence, I am most interested in trends in which SFH is appearing to "slow down".
- The mother's weight is affected by many factors beyond the size of the fetus. During pregnancy, women are typically advised to gain 3-4 times the eventual birthweight of their baby, and issues with fluid retention can cause rapid weight change. Hence, maternal weight is not expected to be a good predictor of IUGR.
- Any observations of changes in SFH ought to be interpreted in light of maternal risk factors for delivering babies that are SGA. These would include smoking status, nutritional status, vascular disease, age, etc.
- Sex of babies does not greatly affect the birthweight except at the highest percentiles. Furthermore, unless there are to be screening ultrasound to determine GA, the sex of the baby will not be known prior to birth. Hence, the applicability of a prognostic model involving fetal sex will depend on availability of at least one ultrasound.

I constructed three variables based on SFH measurements made prior to 30 weeks:

- the lowest ratio of SFH to estimated GA (this is based on the idea that SFH is approximately equal to GA), and
- the rate of change of SFH between two successive visits (change in SFH divided by number of weeks between visits).
- the rate of change of SFH using all measurements between EGA 20 to 30 weeks (least squares slope). Use of this variable in clinical practice presumes that all referrals to high risk clinics would occur at week 30 and that we would have all prior data available for our rule.

The following table displays descriptive statistics (mean, sd, min-max, n) for the minimal rate of change of SFH estimated for each woman up to and including 30 weeks GA. It should be noted that many women did not have two prenatal visits during this period, and thus the sample sizes are lessened. It can be seen that the rate of SFH increase is less in SGA and in smokers, though the pattern in combinations of smoking status and SGA is somewhat confusing. (Data based on comparing the absolute SFH measurements is similar, but a little less pronounced and is not shown.)

| | sga.0 | sga.1 | sga.ALL |
|---|---|---|---|

| | | | |
|---|---|---|---|
| smoker.FALSE | 0.8931 (0.4339, -1.300-3.550; n=397.0) | 0.7680 (0.5282, -1.500-2.000; n= 54.0) | 0.8781 (0.4474, -1.500-3.550; n=451.0) |
| smoker.TRUE | 0.8649 (0.5379, -1.400-2.550; n=171.0) | 0.8540 (0.3442, -0.571-1.500; n= 42.0) | 0.8628 (0.5049, -1.400-2.550; n=213.0) |
| smoker.NA | 0.7433 (0.4082, 0.280-1.050; n= 3.0) | 0.7250 ( NA, 0.725-0.725; n= 1.0) | 0.7388 (0.3334, 0.280-1.050; n= 4.0) |
| smoker.NotNA | 0.8846 (0.4673, -1.400-3.550; n=568.0) | 0.8056 (0.4567, -1.500-2.000; n= 96.0) | 0.8732 (0.4663, -1.500-3.550; n=664.0) |
| smoker.ALL | 0.8839 (0.4668, -1.400-3.550; n=571.0) | 0.8048 (0.4544, -1.500-2.000; n= 97.0) | 0.8724 (0.4655, -1.500-3.550; n=668.0) |

Not presented here are similar analyses that could be done for LBW or pre-term birth.

Similar variables could be defined for changes in maternal weight, with similar descriptive analyses for SGA, LBW, and pre-term.

**Associations with SGA Based on Data at 30 Weeks GA or Before**

The vast majority of you fit logistic regression (OR) models, but RR or RD could also be used. A basic strategy I would have expected to see is an unadjusted analysis of the adverse pregnancy outcomes according to one or more of the three measures of SFH (or weight) changes. Then, I would have adjusted for the maternal characteristics of smoking, age, height, bmi, and parity.

I note that I am not thinking of "confounding" as much in this setting, because the SFH is in some sense a mediator of the other effects. That is, our SFH measurements are just trying to predict fetal growth. If there is IUGR, we probably believe that some risk factor (e.g., smoking) causes slower growth of the fetus, which in turn causes us to notice a fetus that is smaller than usual, which in turn causes a baby to be smaller at birth. This does not really fit with our "confounding" paradigm. Instead, we are adjusting for the other variables, because we are trying to find a diagnostic criterion that is considering women who have underlying pre-pregnancy risk factor and precisely identifying the subset who go on to have SGA, LBW, or pre-term delivery. That is, we are looking for some marker that adds precision to the information we would have had prior to pregnancy.

In answering the question about associations, we would expect to report and estimate, a CI, and a p-value. Interpretation of the estimates is of course key, and as I comment below, it is best if we compare relevant populations.

**Associations with SGA Based on Data at 30 Weeks GA or Before**

Even if the SFH growth rate parameters are statistically significant, that does not mean they are valuable predictively.

Simple measures that we can consider to judge this in a preliminary fashion include:

- The $R^2$. While I am not a fan of using correlation to compare results across different analyses, this is exactly the setting that the $R^2$ is designed for. It tells you whether the available predictors separate the observation into highly homogeneous groups. Certainly, a model that only "explains" 10-20% of the variability is unlikely to be a good predictive model. In this problem, we have a relevant population for

our eventual use of the screening tool, so our estimate of $R^2$ is highly relevant. (In many other data analyses, we have over- or under-sampled the population we care about, and the $R^2$ may under- or over-state the discriminatory capability of our model.)

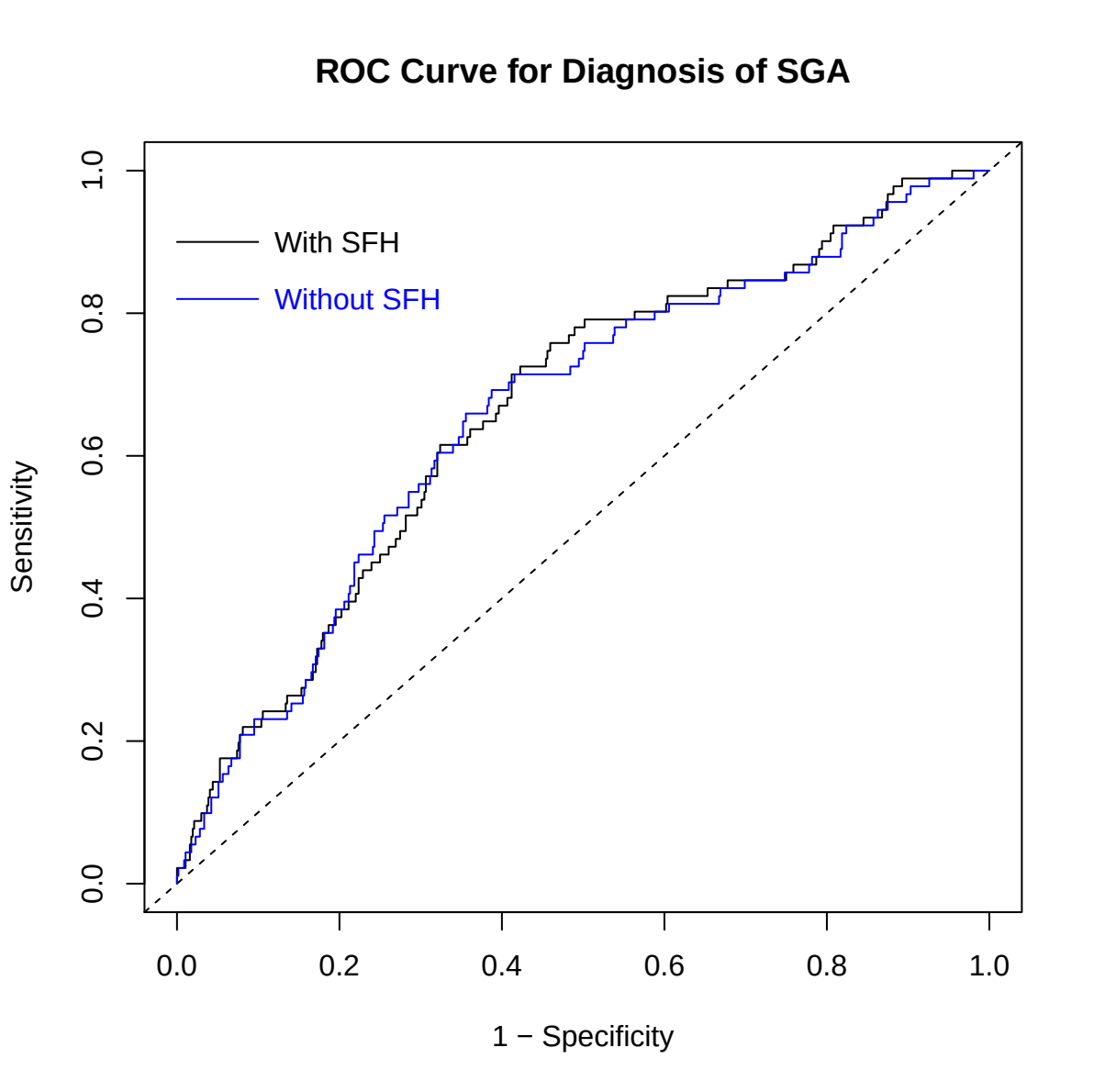
- An ROC curve. This compares the "true positive rate" (sensitivity) to the "false positive rate" (1 – specificity). In judging these values, we also should take into account the prevalence of the condition. For instance:
  - In an ROC curve analysis we could consider a threshold that correctly identifies half the high risk pregnancy (so a sensitivity of 50%). (I chose this arbitrarily—if you prefer something else, use it.)
  - In looking at the ROC curve, this corresponds to a specificity of about 75% (so false positive rate is 25%).
  - In our population based sample (we have data from the exact sort of clinic for which we hope to use the screening tool), there is roughly a 6:1 ratio of low risk to high risk pregnancies. Of the 100 high risk pregnancies, we would find 50% or 50. Of the 600 low risk pregnancies, we would find 25% or 150. Hence, we would use this model to refer almost one-third of the women to the high risk clinic, and three-fourths of those referred women would truly be low risk.
  - To make matters worse, we missed half the high-risk women.

Of course, we did not do proper validation of our model estimates, so the above analysis is too optimistic.

What makes it even worse is the fact that we have nearly the same predictive capability (ROC curve or $R^2$) when using only the pre-pregancy risk factors as we do when we add the variables derived from the SFH or maternal weight measurements.

On the next page, I display ROC curves derived from an RD regression of the SGR outcomes. Similar results were obtained with RR and OR regressions. I display a curve that uses only the pre-pregnancy variables, and I overlay a curve that uses the SFH derived predictors.

It does not take anything fancy to convince me that there is little advantage to be gained by more detailed predictive models.


ROC Curve for Diagnosis of SGA
With SFH
Without SFH
Sensitivity
1 − Specificity
0.0
0.2
0.4
0.6
0.8
1.0

**General Comments on the Reports**

1. The summary is extremely important. Not only does it communicate the key findings of your analysis, but it also helps structure your thinking. Put as much information in it as you can without making it way too long.
    a. Start with the overall goal and specific aims
    b. Give some idea of the amount of information you have available, including
        i. how the data was gathered (characterize it as best you can, trying to think of any biases)
        ii. available measurements
        iii. sample sizes overall and as reduced by missing or irrelevant data
    c. Tell the method of analysis at a very high level (e.g., logistic regression modeling the continuous POI and adjusting for….)
    d. Give results. This includes an estimate, a CI, and a P value. (Yes, yes, there are some analysis methods that do not lend themselves to estimates, but these should be avoided whenever possible.)
        i. I believe there is a special place in hell for people who only report "No significant effect was found". I highlight this in the description of how to write a report. Yet many of you ignore that. Perhaps your advisers tell you to do otherwise. Certainly the literature is full of bad examples. But this time it is easy: I am right and they are wrong. And the goal of a SPH is to make the future better.
    e. Give your top level conclusions.
2. Materials and methods. Explain complicated statistical models. Quite a few of you invoked ROC curves without ever really interpreting them. I do not think AUC helps us at all in this problem, yet it was provided with no real explanation. The ROC curve is useful here, but your reader would be no better at interpreting it than the words you provide.
3. Description of the sampling scheme and types of data that is available. In this problem, a key issue was the description of your derived variables. I was very pleased to see that several reports gave exactly the information that was necessary.
4. Descriptive statistics.
    a. Start with description of the sampling scheme: sample sizes within strata, etc.
    b. When writing to a collaborator, do as many things as you can to identify problems in the data (e.g., one baby was declared SGA even though it was 3780g at birth).
    c. Present descriptive statistics comparing the major POI groups with respect to other covariates. We want to know how else they differ.
    d. Present descriptive statistics that are straightforward estimates of your association. Look at my table above that gave SFH by EGA for the non-SGA and SGA groups. This really tells us everything: IUGR is manifest starting at about 27 weeks—sort of too late for our measures made between 20-30 weeks to do much good.
5. Inferential modeling. Give tables, but give point estimates, CI and p-values in the text for the major analyses.
    a. When you give measure of association (e.g, OR or RD), try to base them on group comparisons that represent about 1 SD of the variation in the sample
    b. E.g., when talking about age of people, do not describe the effect based on a 1 nanosecond difference in age, but also don't describe the effect based on a millennium difference in age

   c. The corollary in this dataset, is that I find the comparisons most of you made about a 1 cm/week difference in SFH growth rates to be so extreme as to be not useful.
6. Discussion: Most of you talked about how further research was necessary for a predictive model based on SFH and maternal weight. I am completely unconvinced. Certainly there is enough information in this data to tell us it is going to be impractical. Don't be afraid of making that clear.