Statistical Considerations in Basic Science Sex Inclusive Research

Suppose an investigator wants to evaluate the effect of a treatment compared to control in a basic science setting involving cell lines. The investigator plans to measure a continuous outcome variable Z in the cells. In what follows, 'female' and 'male' refer to chromosomal female (XX) and chromosomal male (XY), respectively.

Five scenarios pertaining to sex of the cells to be used in the experiment are summarized in the table and described below. These scenarios will impact how the investigator will plan the experiment, including sample size calculations. These scenarios will also impact statistical analyses of the data once they are collected.

Scenario	Is there strong biological justification to study cells of only one sex?	In the control condition, is the mean of Z the same for males and females?	Does treatment affect Z in the same way for females and males?	Is sex a variable that should be included in experiment planning and statistical analyses?
1	Yes	-	-	No
2	No	Yes	Yes	Possibly
3	No	No	Yes	Yes
4	No	Possibly	No, but not of primary interest	Yes
5	No	Possibly	No, and of primary interest	Yes

Scenario 1: There is strong biological justification to study cells of only one sex. All experiments will be carried out only female cells or male cells.

In this scenario, sex would not be a variable included in experiment planning or statistical analyses.

When planning the experiment, sample size calculations can be performed without consideration of the sex variable. The biological justification to study cells of only one sex should be included in grant applications.

When analyzing collected data, sex is not a variable that will need to be considered. When reporting results, biological justification to study only one sex should be reported.

Scenario 2: There is biological justification to study cells of both sexes. The outcome variable Z has the same mean level in cells of both sexes in the control condition, and the investigator anticipates that treatment will cause the same change in Z in cells of both sexes.

In this case, sex may be a variable to be included in experiment planning and statistical analyses.

When planning the experiment, if it can be reasonably assumed (either based on preliminary data or published literature or both) that mean levels of Z in the control condition are similar for both sexes and that treatment has the same effect in both sexes, then sample size calculations will be no different than calculations made in Setting 1. If it is possible that mean levels of Z and/or the effect of treatment on Z differ by sex, then Scenarios 3, 4 or 5 may be relevant.

When analyzing collected data, the assumption that Z has the same mean in cells of both sexes in the control condition should be examined using descriptive statistics. The similarity of the treatment's effect on Z for cells of both sexes should also be examined using descriptive statistics. If assumptions are met, it may be appropriate to analyze data from the female and male cells together, without statistical control for sex. If either or both of the assumptions are not met, then analysis approaches for Scenarios 3, 4 or 5 are warranted.

Scenario 3: There is biological justification to study cells of both sexes. The outcome variable Z has a different mean level in female cells and male cells in the control condition. While the control means in females and males are different, the investigator anticipates that treatment will cause the same change in Z for cells of both sexes. Stated another way, the investigator anticipates that differences in Z for treatment versus control will be the same for female cells and male cells.

In this case, sex is a variable that should be included in experiment planning and statistical analyses.

When planning the experiment, sample size requirements may be slightly higher than in studies of one sex only. Actual requirements will depend on the anticipated extent to which the mean of Z in the control condition differs for female and male cells.

When analyzing collected data, sex should be included as a 'covariate' in all analyses, including regression. The similarity of the treatment's effect on Z for cells of both sexes should be examined using descriptive statistics. It may be reasonable to report means of Z in the control condition and after treatment separately for female and male cells.

Scenario 4: There is biological justification to study cells of both sexes. The outcome variable Z may or may not have different mean levels in female cells and male cells in the control condition. The investigator anticipates that treatment will affect the outcome Z differently for female and male cells. Stated another way, the investigator anticipates that the difference in Z for treatment versus control will be different for female cells and male cells. While a difference in treatment effect depending on sex may be anticipated, formal statistical testing of this difference is not of primary interest to the investigator.

In this case, sex is a variable that should be included in experiment planning and statistical analyses.

When planning the experiment, sample size requirements will be higher than in studies of one sex only. Actual requirements will depend on the extent to which the mean of Z in the control condition and/or the treatment's effect on Z are anticipated to differ for female and male cells.

When analyzing collected data, sex should be included as a 'covariate' in all analyses, including regression. An 'interaction term' between treatment and sex should also be included to account for the difference in treatment effect according to sex. It is important to report means of Z in the control condition and after treatment separately for female and male cells.

Setting 5: There is biological justification to study cells of both sexes. The outcome variable Z may or may not have different mean levels in female cells and male cells in the control condition. The investigator anticipates that treatment will affect the outcome Z differently for female and male cells. Stated another way, the investigator anticipates that the difference in Z for treatment versus control will be different for female cells and male cells. Formal statistical testing of the difference in treatment effect for female and male cells is of primary interest to the investigator.

In this case, sex is a variable that should be included in experiment planning and statistical analyses.

When planning the experiment, sample size requirements will be higher than in studies of one sex only. Specifically, the study should have adequate sample size to test whether the interaction term between treatment and sex is statistically significantly different from zero. Actual requirements will depend on the extent to which the mean of Z in the control condition and/or the treatment's effect on Z are anticipated to differ for female and male cells.

When analyzing collected data, sex should be included as a 'covariate' in all analyses, including regression. An 'interaction term' between treatment and sex should also be included to account for the difference in treatment effect according to sex. Statistical testing of this interaction term will be of primary interest. It is important to report means of Z in the control condition and after treatment separately for female and male cells.