2. Study Design and Data Collection

The inferences we may draw from the data depend crucially on the study's design and the data collection process. We will describe the study design by answering the following questions:

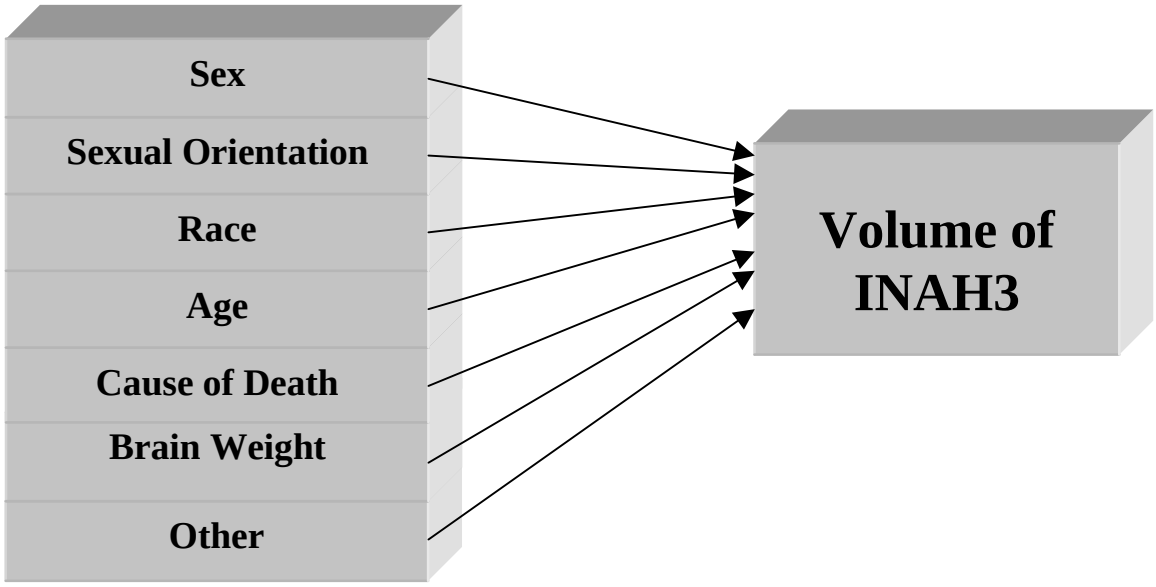
2.1 What factors might affect the volumes of INAH3?
2.2 What is the purpose of the study?
2.3 What is the scope of inferences in the study?

Before we will discuss the scope of inferences in the study, we have to identify important variables that might affect volume of INAH3.

2.1 What factors might affect the volumes of INAH3?

According to pathologists, there are several factors that might affect the size of INAH3. The most important are sex, sexual orientation, race, cause of death (AIDS), age, brain weight, hospital of origin, length of time between death and autopsy. In general, it would be difficult or even impossible to consider all variables affecting the volume. Some variables may not be recognized or measured.

The variables that might affect volumes of INAH3 are visualized in the following diagram.



According to biological research, sex, sexual orientation, and brain weight are recognized as the most important variables affecting the size of INAH3. In particular, it has been shown in the recent years that INAH3 is more than twice as large in the heterosexual men as in the women. As the subjects in the case study had similar ages, we expect that the age factor does not contribute significantly to the observed differences.

2.2 What is the purpose of the study?

The goal of the study is to examine the relationship between sexual orientation and volume of INAH3. Specifically, the researcher hypothesized that INAH3 is large in individuals sexually oriented toward women (heterosexual men and homosexual women) and small in individuals sexually oriented toward men (heterosexual women and homosexual men).

The five groups in the study are some combinations of the three categories: sex (male or female), sexual orientation (heterosexual and homosexual), and cause of death (AIDS and non-AIDS). Including cause of death is a consequence of the fact that the volumes of INAH3 were obtained only for the homosexuals who died of AIDS. Moreover, there is also a possibility that the size of INAH3 is affected by AIDS or its complications. The response variable is the volume of INAH3.

2.3 What is the scope of inferences in the study?

The five groups in the study are based on some combinations of sex, sexual orientation, and cause of death of the 41 subjects at autopsy. As none of the factors was decided by the investigator for each of the 41 subjects, the case study is an example of an observational study. In other words, allocation of the subjects to the five groups was not determined by any chance mechanism.

As the study is an observational study, we are not able to draw any causal conclusions from the statistical analysis alone. It is possible that some confounding variables are responsible for the disparity in the volumes of INAH3 between heterosexual and homosexual males. For example, as all homosexual males in the study died of AIDS, there is the possibility that the small size of INAH3 in the homosexual men is the result of AIDS or its complications and is not related to the men's sexual orientation.

Until tissue from homosexual men dying of other causes becomes available, the possibility cannot be rigorously excluded. Therefore, the results will not allow us to decide if the size of INAH3 in an individual is the cause or consequence of that individual's sexual orientation.

Notice that the 41 subjects at autopsy in the five groups were not selected from any well-defined populations. Therefore, the observed pattern cannot be inferred to hold in some general populations, for example the population of all homosexual men unless we assume that the homosexual men in the study are representative of the population. However, all homosexual men in the study had AIDS. There is a possibility that AIDS patients constitute an unrepresentative subset of gay men.