### Addressing the Hierarchical Data Structure

Because the study is hierarchical, an overview of the connection between the files is provided. Data can potentially be aggregated from one level into another level that can potentially create a disclosure risk. [*Provide an overview of the connection between the files. For example, the race or gender of students of a school can be aggregated to the school level and provide potentially reliable information on the identity of a school when other factors are considered. It is important to mask the cluster identity, either because there is a questionnaire associated with the cluster (such as a principal questionnaire for a sampled school) or there is risk associated with the identity of the cluster (such as prisons in a prison survey), then the risk assessment and data swapping process would be conducted at each level.*

*Whether or not the data are hierarchical or not, some studies share/include data from other studies. For example, natality data from an outside agency may be shared with an early childhood education study. These data must be evaluated when determining disclosure risk for both studies as this could compromise data from both agencies.*

*A description of all levels and how it impacts processing should be described here. For example, each level of data (I.e. school, teacher, parent, and student) is required to undergo controlled random swapping. Assessment files are generally not required to be included in any swapping procedures. Each level can be done separately and independent due to the swapping variables being considered. At each level, the swapping is often processed five times (using different random seeds) before selecting the best swapped dataset based on the DataSwap impact measures, and according to the IES DRB guidelines. The best run in terms of the global utility measures will be selected to serve as the official swapping file. Prior to swapping, the InitialRisk software can help facilitate the risk assessment and can inform the swapping process at each level.*

*If any school level variable can be derived from student or teacher level variables, they can (if they can match up with a publicly available file) potentially be included in the probabilistic matching procedures.]*