## 1.9 Data Cleaning Procedures

[*Describe the procedures implemented for data cleaning*.

* *Are any data out-of-range? How are they handled?*
* *Are any valid data outliers? If so – are they top- and bottom-coded?*
* *Are data keyed and verified?*
* *Are data scanned?*
* *Are they imported electronically?*
* *Are they edited during collection or post-collection?*
* *Are any data imputed and if so, are imputation flags included on the file?*

*Why does this matter? It is helpful to know if a study has highly reliable data or if a lot of noise exists. This may help the DRB chair factor in which swapping variables and which swapping rates make the most sense. The cleaning procedures should precede the SDC procedures. If data are imputed, that is important to describe – are imputation flags used and included on the file? If not, can imputed data be factored into the risk assessment?* ]