**Appendix 2**

*Example of Missing Data Patterns*



Next to different missing data mechanisms (i.e., MCAR, MAR or MNAR), two general missing data patterns can be distinguished in single-level data, namely **univariate** missing data and **multivariate** missing data. We illustrate these missing data patterns by means of a simplified hypothetical example, consisting of a sample of 3 children. For each child, age (at first consultation), gender, ethnical back-ground and nutritional status (in 3 time points) were registered. A missing data pattern is said to be **univariate** if only one variable of the dataset has missing observations and the rest of the variables is completely observed, see child 1 in the example above. A missing data pattern is said to be **multivariate** if data is missing in more than one variable, as is the case for children 2 and 3. In case of multivariate missing data, several situations may occur. In some cases, such as for child 2, only the first assessment of nutritional status is observed and the subsequent ones not as a result of the child dropping out of the study or being lost to follow-up (e.g., changed schools). However, in other situations, such as for child 3, there are intermittently observed values. In other words, a missing value is preceded and/or followed by an observed value (e.g., the child missed school the day of the second measurement). This is referred to as a **non-monotone** missing data pattern.

*Example of Missing Data Patterns in Multilevel Data*



Individuals in epidemiological studies are often clustered within units, for example, different locations of one healthcare organization or hospital. This clustered or dependent nature of individual-level observations is referred to as **multilevel data**. Clustering is also considered present when repeated assessments are done within an individual. It is important to account for this clustering of the data in the imputation process with patients from the same unit being more similar than patients from different units. Missing data can be present at both the individual and the unit-level as described below.

In multilevel data, additional missing data patterns may occur. We illustrate this using a hypothetical example, in which three schools are included, all of which contain three children (that is, children are clustered within schools). Thus, the children are considered the **level-1 units**, whereas schools are the so-called **level-2 units** or clusters. Similar to single-level data, in multilevel data missing observations can be present in outcome variables and individual level covariate (s). In addition to single-level data missing data, missing observations can also be present in level-2 covariate(s) (i.e., cluster level Covariate(s)s in multilevel data). For instance, some clusters may not register a specific characteristic (e.g., ethnical background in school cluster B), resulting in missing observationsfor all individuals in a cluster. Finally, for two schools the type of school is known (public/private), but for the third school this is unknown, also resulting in missing observations at the cluster level (i.e., level 2). Thus, in comparison with single-level data, complexity is added when missing observations in multilevel data are present at different levels.