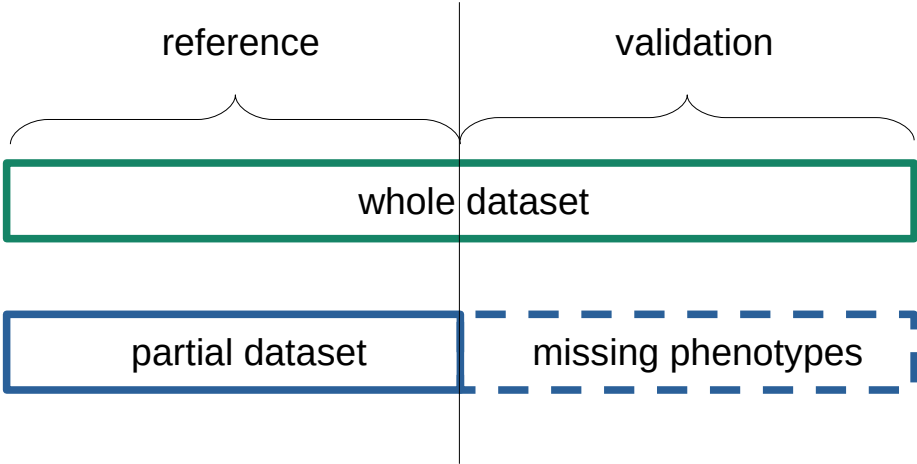


reference

validation



The diagram illustrates dataset partitioning across two rows and two columns. A vertical line separates the 'reference' column on the left from the 'validation' column on the right. In the top row, a green-outlined box labeled 'whole dataset' spans both columns, with curly braces above it indicating its extent. In the bottom row, a blue-outlined box labeled 'partial dataset' is located in the reference column, and a dashed blue-outlined box labeled 'missing phenotypes' is located in the validation column.

whole dataset

partial dataset

missing phenotypes