

11.1 Reliability of Patient Data

High quality data collection is imperative to the quality and accuracy of analytics results, irrespective of the terminology used. Whether the focus is decision support, business intelligence, research or a mixture of all three - data quality is critical. High quality information is not the consequence of collecting as much data as possible. Instead, it is the product of intentionality and process design.

The factors that may impact the quality of patient data include:

- The design of user interfaces used to capture data

Clinical user interfaces should be designed to make it as easy as possible to find the most appropriate code, and as difficult as possible to enter the wrong code. There are a variety of ways to improve the ease and effectiveness of data entry using SNOMED CT – such as searching over all synonyms, confirming the selected concept using the preferred term or fully specified name, ordering value lists effectively using an ordered reference set, searching using navigation hierarchies, and constraining data entry using subsets.

 These techniques can also help to reduce data entry errors by prohibiting invalid input, helping the user to understand the correct meaning of the code selected, and ordering value lists in a clinically safe order (e.g. ordering medications by strength, rather than alphabetically).

- Use of diagnostic criteria to standardize data capture

Diagnostic criteria and their application tends to vary widely according to care setting, patient status and healthcare professional. The consistent ascertainment and recording of even common diagnoses, such as asthma and myocardial infarction is often non-trivial. High quality prospective research studies require that diagnostic criteria for the condition being studied are understood, rigorously applied and accurately documented. In routine clinical practice doing this for potentially thousands of diagnoses in dozens of care settings is normally infeasible. Divergence and inconsistencies in criteria for diagnosis capture can undermine the validity of any conclusions which may be drawn from analytics. SNOMED CT mitigates this issue by allowing the query author to choose a reliable aggregating concept from SNOMED CT's extensive content.

- Consistency of data capture with analytics requirements

Pick lists and constraints should be consistent with both clinical data collection needs and analytic requirements and these should never be in conflict. The presence or absence of particular concepts in value sets within different applications can cause data collection to be inconsistent. SNOMED CT mitigates this by allowing the query author to choose a reliable aggregating concept.

- Loss of meaning during data transformations

Clinical data often undergoes a number of structural transformations and code mappings prior to data analytics being performed, during the process of preparing the data for messaging and/or loading into a data warehouse. In each of these transformations, care must be taken to ensure that the quality of the process is high, and that there is no incremental shift in the clinical meaning of the data. For example, mapping local codes to an alternate code system using non-equivalence maps (e.g. narrow to broad or broad to narrow) will change the clinical meaning of these codes to some degree. Any changes that effect the clinical meaning of the data may have an impact on the quality of data analytics. SNOMED CT helps to mitigate this by supporting the representation of equivalence maps, which can be used when the use case requires.

 *SNOMED CT Search and Data Entry Guide*, 2014, <http://snomed.org/search>.