12.1.1 Data Migration Workbench (UK)

The workbench is produced by the UK Terminology Centre (UKTC) as a DRAFT ‘proof of principle’ product. It demonstrates advanced functionalities, leveraging SNOMED CT as a sophisticated reference terminology together with the mappings across the current five NHS terminologies or classifications. It is not currently intended to develop this as a fully supported product.

The UK Terminology Centre Data Migration Workbench (DMWB) is designed to support the NHS Primary Care Summary Care Record, Primary Care Systems of Choice and Data Migration programs. This tool demonstrates some of the properties and advanced uses of the data migration and mapping products published by the UKTC and the terminologies and classifications that they link.

The workbench uses SNOMED CT to perform novel and sophisticated analyses of patient data. It has immediate 'off the shelf' international utility despite the inclusion of the UK-only terminologies within the standard tool distribution.

The software contains SNOMED CT, Read Codes Version 2 and CTV3, maps between these and maps to ICD-10 International Edition (UK map not the same as the SNOMED International one) and OPCS Classification of Interventions and Procedures (OPCS-4). The Workbench modules support:

1. Searching and browsing the hosted code systems;
2. Viewing maps between the hosted code systems;
3. Authoring analytics subsets (i.e. terminology query predicates) and their testing, maintenance and 'translation' between code systems;
4. Electronic Patient Record (EPR) data quality analysis and data repair; and
5. EPR reporting and case mix analysis.

Queries Tool

The Queries Tool offers advanced functionality for authoring, maintaining and testing query code sets (called ‘clusters’) or subset definitions within any of the supported terminologies or classifications. One major application is to produce query sets which will return comparable results from patient records encoded with different code systems. To assist with this, the tool translates subset definitions expressed using one terminology into subset definitions expressed using another, allowing refinement by manual editing (see figure 12.1.1-0 and figure 12.1.1-0).

Figure 12.1.1-0: DMWB Queries Tool building a SNOMED CT asthma subset
Electronic Patient Record Data

The EPR Data Tool provides an environment for loading and analyzing patient data, either to design query specifications or as part of data quality and case mix analytics. The analytics functions include detection and management of coding data quality issues such as:

- Records with inactive SNOMED CT codes;
- Records with codes from inappropriate SNOMED CT hierarchies e.g. a diagnosis recorded using a concept from the substance hierarchy (e.g. 419442005 |ethyl alcohol|) rather than a code from the disorder hierarchy (e.g. 25702006 |alcohol intoxication| or 7200002 |alcoholism|).

The tool also enables the rapid repair of such data by substituting inappropriate codes with more appropriate ones. This service is performed in an offline reporting environment.

Analytics

The workbench data analytics tool runs cluster queries, combined with demographic data, to perform clinically valuable case finding, case mix and caseload analysis.

The 'Overview' Report module includes:

- Basic demographics (population age, sex, ethnicity);
- Analyses of episodes with a SNOMED CT code;
- Counts by SNOMED CT supercategory;
- List of the 100 most frequently used individual SNOMED CT codes; and
- List of the 15 most common SNOMED CT codes for each age cohort.

The Trends module analyzes the frequency with which individual SNOMED CT codes are used in the EPR instance data, looking for those whose recording frequency has changed over the course of the data collection period.

The Induce module performs a more sophisticated analysis of case mix and caseload trends within a clinical department. Instead of returning the most frequently used individual codes, the Induce module attempts to identify the most frequently used types of codes. For example, an emergency department may use roughly 500 different SNOMED CT codes for a laceration in a particular anatomical location. While none of the site-specific codes may appear in a list of most common codes, the descendants of 312608009 |laceration| may collectively account for a significant part of the department's workload.

The Graphs tool performs fundamentally the same query and search operations, but generates graphs based on the patients or episodes identified, showing e.g. the age:sex distribution of patients in a defined case mix cohort, or the changing incidence of one or more specified clinical phenomena (e.g. disease presentation, or procedure performed) by year, quarter, month or day of the week. These graphs can be copied into documents.

https://isd.hscic.gov.uk/trud3/user/guest/group/0/pack/8