CASE STUDY
Future Healthcare Journal 2017 Vol 4, No 2: 126–30

Implementation of SNOMED CT in an online clinical database

Authors: Mark Wardle[A] and Andy Spencer[B]

ABSTRACT

A number of strategies have been published to accelerate the use of electronic health records in caring for patients across the UK. These visions of 'eHealth' have a common requirement for robust interoperability between different systems with the use of appropriate information and data standards. SNOMED CT, a comprehensive terminology that NHS England intends to adopt across all care settings by 2020, is a key component of these standards but there is currently limited experience in its use in live clinical settings. Within NHS Wales, an electronic patient record system has been developed since 2009 with a focus on a core generic clinical information model built using SNOMED CT. Our experience is that SNOMED CT is a usable and clinician-friendly terminology but that its size and scope must be considered during implementation.

KEYWORDS: Clinical databases, clinical terminology, medical informatics, SNOMED-CT

Introduction

The NHS England *Five-Year Forward View* is that the 'NHS should go paperless by 2020';[1] NHS Scotland's eHealth strategy is predicated on 'information sharing and communications that facilitate integrated health and social care across all settings from the patient's home to the hospital';[2] NHS Wales' *Informed health and care* 5-year strategy focuses on creating an open platform to support interoperability between systems;[3] the Northern Ireland *eHealth and care* strategy focuses on greater standardisation with the use of structured information to deliver a 'fully integrated electronic patient record'.[4] In all of these visions, there is a common requirement for robust interoperability between different systems with the use of appropriate information and data standards.

Since 2009, an all-Wales neurological electronic patient record system has been developed to provide clinicians with a simple way to record diagnostic, treatment and interventional information using a web-based portal. The database, underpinned by Systematized Nomenclature of Medicine Classification Terms (SNOMED CT), has proven clinical benefits, including supporting

- direct patient care by making information available immediately to staff at the point of care
- clinical governance and service provision by making aggregated data immediately available in live reporting
- clinical research with a focus on generic and disease-specific patient outcomes.

An understanding of what has been achieved through the online neurology database and the process by which this came about will provide a valuable insight into how SNOMED CT might be implemented across the UK together with the resulting benefits.

Concept behind the database build

The database was developed to address a number of important clinical requirements in our neurological services that could not be met using the existing paper and electronic systems that were in place in 2009 when the development started. Answers to fundamental questions about individual patients or clinical services, such as numbers of patients seen with particular diagnoses, their treatments and outcomes, were beyond reach. Information on numbers of outpatient appointments could be requested but only anecdotal data relating to why a patient was being seen were available. There was no information about what had been done or how the patient had fared.

The organisation of neurology services in Wales required an approach that was not limited by organisational boundaries. Therefore, one important goal was to ensure that the database was centred on the patient and accessible to all staff involved in the care of the patient. This requirement needed a more sophisticated access control mechanism than most organisation-centric systems, which usually limit access to their clinical staff. As such, access control was based on logically organised clinical services with access determined by a carefully curated list of team members on a per service level basis.

Recognising that any individual database must not be a standalone system but instead needs to integrate and interoperate with a number of other current and future systems, there was an early focus on data standards and how the database would support the safe exchange of information. Furthermore, an information-first approach would ensure future extensibility and system longevity as different user

Authors: [A]consultant neurologist, Cardiff and Vale University Health Board, Cardiff, UK; [B]national clinical lead for hospital specialties, Health and Social Care Information Centre, Leeds, UK

interfaces, such as web-based or tablet devices, could be used as a front-end to the core underlying information model.

It is common to think of databases in terms of appearance and function. Instead, this development focused on important high-level clinical information requirements and the development of a core information model that would underpin those requirements. This can be achieved by examining high-level functional clinical requirements and decomposing them into a set of core information requirements. For example, one high-level requirement was to 'support complex epidemiological studies of neurological disease deriving automatic incidence and point-prevalence figures for a range of disorders'. This requirement needed a list of patients, their current and past addresses, a way of mapping addresses into geographical regions and a record of diagnostic information, including the date of diagnosis. To map addresses to geographical regions, we used the Health and Social Care Information Centre's (now NHS Digital) organisation data service – which includes data to map postcodes into health regions. The remaining issue was the recording of diagnostic information.

In developing the information model to underpin the clinical information system, there were two important decisions to be made. Firstly, appropriate terminology – a 'body of terms' much like words in a dictionary – is required. There are a number of different clinical terminologies available, such as Read codes (as used in primary care), ICD-10 (a disease classification used in clinical coding) and SNOMED CT (developed from an amalgamation of Read codes and an American pathology database). While these terms have meaning when used in isolation (eg 'myocardial infarction'), it is only when these terms are combined together in a logical way as part of a larger data model that true meaning can be understood. It is analogous to definitions for individual words in a dictionary, but true expression results when these words are combined into sentences and paragraphs. As such, one must also consider the data model that will provide a structure in which terms can sit and while there are a number of these available (eg open electronic health record archetypes[5] and Health Level Seven International[6]), there is no single standard at the time of writing.

Recording clinical terms

Could free text be used to record clinical terms? Many small departmental databases use free text, but free text is inadvisable as users can make errors in data entry and such errors make subsequent analysis extremely difficult. Another option is to develop a list of common diagnoses and allow users to choose from the list. A defined code or the text is then stored. However, such lists limit users to diagnoses that have been included in the list unless free-text entry is also permitted. Inclusion of free-text and own-coded information will make searching and subsequent interoperability with other systems difficult or impossible.

Read codes are an established system of recording clinical information. They are based on a hierarchical system of codes with each level becoming more specific. Searching for a code beginning with F will show all terms relating to diseases of the nervous system with Parkinson's disease represented by 'F12z' and motor neurone disease by 'F152z'. Its simplicity is a disadvantage and many neurological diagnoses are not represented, making it less useful for hospital medicine. In addition, should bacterial meningitis be in the neurological disease or the infectious and parasitic diseases hierarchy? In Read, such issues have frequently resulted in duplicate codes.

SNOMED CT is a very large and comprehensive terminology. Importantly, terms can be linked by multiple relationships to other terms. This makes it possible for software to determine that multiple sclerosis is a disease defined by 'demyelination' of the 'central nervous system'. When implemented properly, SNOMED CT enables software to make 'intelligent' decisions about what to show, what data to request and what forms to present based on the diagnoses entered. For example, the database would 'know' that a patient had epilepsy if they were given a diagnosis of 'juvenile myoclonic epilepsy' or 'frontal lobe epilepsy' or any of the 143 (at the time of writing) other terms that are equivalent to a diagnosis of epilepsy. Thus, a command to 'send an alert when a patient, belonging to a particular consultant, with motor neurone disease loses 5% of their body mass compared to their baseline at diagnosis' can be implemented easily. SNOMED CT allows the underlying logic to simply ask whether the patient has a type of 'motor neurone disease' and this would automatically include all patients with related diagnoses, such as 'primary lateral sclerosis' and 'pseudobulbar palsy'.

SNOMED CT is not confined to diagnostic and procedural information. There are hierarchies covering a wide range of medical terminology, including anatomical structures, pathology, occupations and ethnic origins. With local extensions, such as the NHS' dictionary of medicines and devices, these codes can be used in any field that needs structured, coded information.

Another advantage is support for synonyms. A distinct clinical concept can and usually has multiple synonyms – for example 'granulomatosis with polyangiitis' was previously known as 'Wegener's granulomatosis'. With synonym support, a user entering an outdated or synonymous term would find the synonym and see it mapped into the new, modern, preferred description of the term.

Within SNOMED CT, clinical terms are 'concepts', synonyms are 'descriptions' and the relationships between concepts are recorded as 'relationships'. While seemingly simple, as relationships themselves are defined by concepts (such as 'IS-A' as in 'motor neurone disease' – 'IS-A' – 'disorder of nervous system'), it means that the relationship tree is infinitely extendable over time.

SNOMED CT is owned by SNOMED International (formerly the International Health Terminology Standards Development Organisation) and is an international terminology, with the UK version managed by the UK Terminology Centre (UKTC) of NHS Digital. A detailed understanding of SNOMED CT can be obtained through the online training resources provided by SNOMED International[7] and access to the online SNOMED CT browser.[8]

Given these benefits, we chose SNOMED CT as the underlying terminology for the neurology database.

Implementing SNOMED CT

It was essential to make SNOMED CT ubiquitous at the information level of the database and as the *lingua franca* both between modules within the system and in interfacing with other systems externally. Many electronic patient records bolt on a terminology service as an extra, as simply a way of showing or searching a list and then recording the term. Instead, this system integrated SNOMED CT fully and allowed multiple methods for users to enter terms without necessarily realising that they are using a terminology service at all. This required us to implement several important items of functionality.

Firstly, in a given data field, a user is able to enter a clinical term. Usually, this is in a pop-up list of common terms for that field. The terms displayed are determined by contextual data derived from the patient, the logged-in clinician and the clinical services to which they are both registered. Therefore, the common terms in an epilepsy clinic setting are different to that in a paediatric clinic. When the required term is not identified, the interface allows the user to expand their search and type in real time much like that available when searching Google. In other words, the search occurs as the user types and the user feels as if they are typing free text but it is instantly mapped to a SNOMED CT term. For most data fields, common terms and subsequent searching is filtered so that only a subset of appropriate terms are searched, such as those representing a clinical diagnosis. In a field for entering country of birth, the search would be limited to terms that represent this concept.

Secondly, the recording of generic and disease-specific outcome measures was implemented to include those reported by patients and those captured by clinicians during the process of care; the Expanded Disability Status Scale (EDSS) in patients with multiple sclerosis and the Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS) in patients with motor neurone disease are two important examples of this. The use of such structured rating scales allows SNOMED CT terms to be derived based on the information entered using the scale. For example, if a patient is recorded as having an EDSS score of 4.0 or less, then the patient can be recorded as having the finding of 'independent walking'. Similarly, users are unaware that they are recording SNOMED CT terms as they click the sites of injections of botulinum toxin during a procedure.

Thirdly, rapid recording of new and changed medications in an outpatient setting was implemented. The user is able to rapidly type out a list and have their text automatically parsed and converted into SNOMED CT terms on the fly without pop-ups or clunky and slow interfaces.

Finally, the system uses the semantic knowledge within SNOMED CT to support clinical decision making. Given a specific term, computer logic can be written that allows 'intelligent' advice to be given.

SNOMED CT was first implemented in the system in 2009. Refinements have resulted in a standalone web service that provides similar functionality to external databases.

Benefits derived from using SNOMED CT

The most important benefit has been an ability to understand patient cohorts. Simply recording accurate clinician-derived diagnostic information has informed our clinical services. We have published large-scale epidemiological work on multiple sclerosis, its incidence and prevalence using our SNOMED CT-driven database.[9] We have prospective longitudinal cohorts of patients with a range of neurological diseases, including multiple sclerosis, Parkinson's disease, motor neurone disease and neuromuscular disease. We have just launched an all-Wales epilepsy register. We now have over 20,000 patients registered in our database.

Another benefit is the ability to compare an individual patient with the whole cohort of similar patients in real time while sitting in an outpatient clinic. By graphing disability over time (Fig 1), we can see that this patient with multiple sclerosis has done better than expected, probably as a result of receiving alemtuzumab early on in the course of the disease. Comparing disease progression or activity supports clinical decision making and allows an informed discussion with a patient, allowing them to see how they are doing relative to their peers.

Flexibility to add functionality to the system is another benefit of using a comprehensive and wide ranging terminology system. If another team of clinicians wished to record their inpatient and outpatient diagnoses then this would simply need some minor reconfiguration of the database, even in an unrelated specialty like nephrology or gastroenterology. For example, the botulinum toxin therapy clinic saw a range of patients from across south Wales, but local clinicians and managers had little idea about how much toxin and which brands were being used. It took 3 hours of development to implement a system to record toxin administration with anatomical sites, the procedure, toxin and complications recorded using SNOMED CT with some terms entered by choosing sites on anatomical drawings. The clinical document produced contains the letter but also all of this structured information; another system that implements SNOMED CT could similarly understand these data and make appropriate inferences.

Future benefits of SNOMED CT

As the database is built using SNOMED CT, interoperability with other systems in Wales is facilitated. For example, the NHS in Wales is developing a service-orientated architecture in which core services, such as demographics and clinical documents, are stored centrally with a flexible architecture designed to support multiple applications. As part of a new Welsh Care Records Service, clinical documents can be stored together with relevant metadata. With SNOMED CT, the neurology service can send documents together with metadata encoded in SNOMED CT and thus letters could be searched for diagnostic or procedural terms automatically as part of a faceted search page.

Pitfalls to be avoided in implementing SNOMED CT

SNOMED CT is a large terminology containing over 1.6 million synonyms. As such, technical implementation is not trivial and requires several optimisations in order to allow fast free-text searching. Indeed, searching using SNOMED International's own web-based browser[8] is frequently slow and would not be suitable in a live clinical environment.

In addition, there are 1,756 synonyms containing the word 'pneumonia'. A naive implementation that shows all of these would not be usable in clinical practice. It is therefore useful to be able to show curated lists in which commonly used terms are shown before allowing a broader search. The neurology database determines these curated lists in real time based on

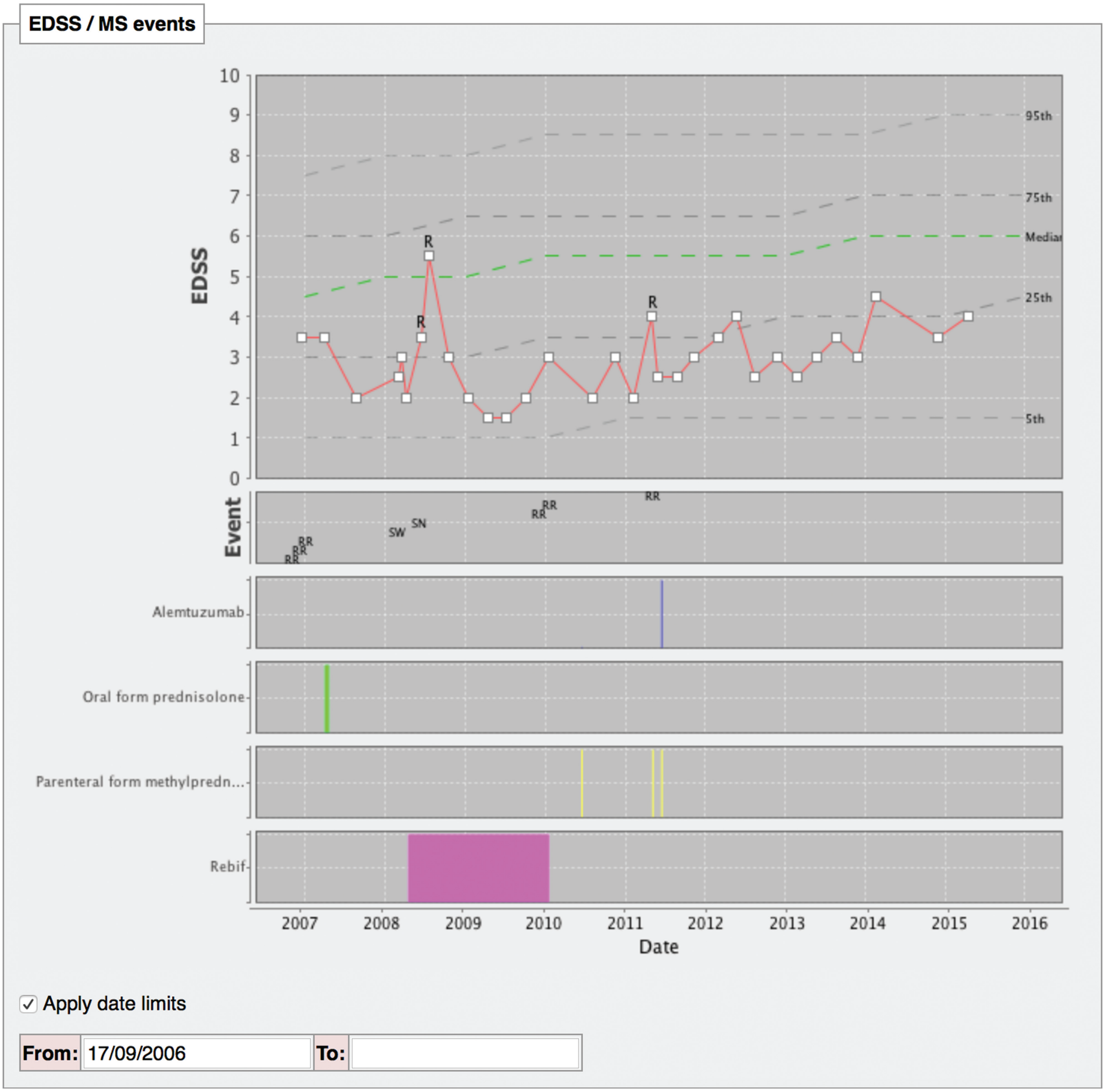

Fig 1. Comparing individual patient performance with the whole cohort over time, in real time. The patient's course (red line) is plotted against progress for all patients. SNOMED CT allows the database to identify other patients with the same disease and to filter and show only disease-modifying treatments rather than all of the treatments received by the patient.
EDSS = Expanded Disability Status Scale; MS = multiple sclerosis; SNOMED CT = Systematized Nomenclature of Medicine Classification Terms

contextual information, such as knowledge about the patient and the user and the clinical services and research projects to which they are registered. In addition, even more broad searches are carefully filtered so that inappropriate terms are not visible. For example, when choosing country of birth, the neurology database limits searching to concepts that reflect that semantic concept.

Finally, SNOMED CT is continually updated and any implementation must, therefore, ensure that there is a robust system to new releases that may result in terms becoming deprecated and new terms being introduced.

Costs of implementation

The content representing SNOMED CT is available free of charge to the NHS via NHS Digital, subject to agreeing to SNOMED International and UKTC licences.[10] However, making SNOMED CT usable in practice – both to enter terms and to understand what a term means semantically – requires a highly-optimised terminology service that can be embedded or used within an application. One of the authors has released a freely available open-source terminology server that can be used in third-party applications.[11]

The neurology database consists of this terminology service embedded within a larger web-based service that is now used by a range of different medical specialties. This requires ongoing maintenance on a server hosted by the organisation.

Conclusions

We recommend an early focus on data modelling and information standards when designing clinical information systems. We think that SNOMED CT is an important core component, providing a facility to allow members of the clinical team to contemporaneously record structured diagnostic and outcome information about their patients. Importantly, implementers should carefully consider how to embed SNOMED CT and not simply view this terminology as a 'bolt-on' for their systems.

Fully integrating the semantic understanding of clinical terms and then making individual and aggregated patient information available at the point of care can support clinical practice. It can also ensure that clinicians gain immediate benefits from recording prospective data.

The use of SNOMED CT allows a system originally designed for use in neurological practice to be used more widely with little change. When used within a robust and open information model, SNOMED CT facilitates interoperability with other clinical systems within the organisation. ■

Conflicts of interest

MW is co-director of a medical software and consultancy company and has released an open-source SNOMED CT terminology server.

Author contributions

MW wrote the initial draft and both authors made subsequent revisions and reviews.

References

1 NHS England. *Five Year Forward View*. London: NHS England, 2014.
2 NHS Scotland. *eHealth strategy 2014–2017*. Edinburgh: The Scottish Government, 2015.
3 NHS Wales. *Informed health and care: a digital health and social care strategy for Wales*. Cardiff: Welsh Government, 2015.
4 Health and Social Care Board. *eHealth and care strategy for Northern Ireland: improving health and wealth through use of information and communication technology*. Belfast: Health and Social Care Board, 2016.
5 open EHR. www.openehr.org [Accessed 31 January 2017].
6 Health Level Seven® International. www.hl7.org [Accessed 31 January 2017].
7 SNOMED International. www.snomed.org/[Accessed 25 April 2017].
8 SNOMED International SNOMED CT Browser. http://browser.ihtsdotools.org [Accessed 31 January 2017].
9 Harding KE, Wardle M, Moore P *et al*. Modelling the natural history of primary progressive multiple sclerosis. *J Neurol Neurosurg Psychiatry* 2015;86:13–9.
10 Health and Social Care Information Centre. UK Terminology Centre. https://isd.hscic.gov.uk/ [Accessed 31 January 2017].
11 SNOMED CT terminology server. https://github.com/wardle/rsterminology [Accessed 25 April 2017].

Address for correspondence: Dr M Wardle, Cardiff and Vale University Health Board, University Hospital Wales, Heath Park, Cardiff CF14 4XW, UK.
Email: mark@wardle.org