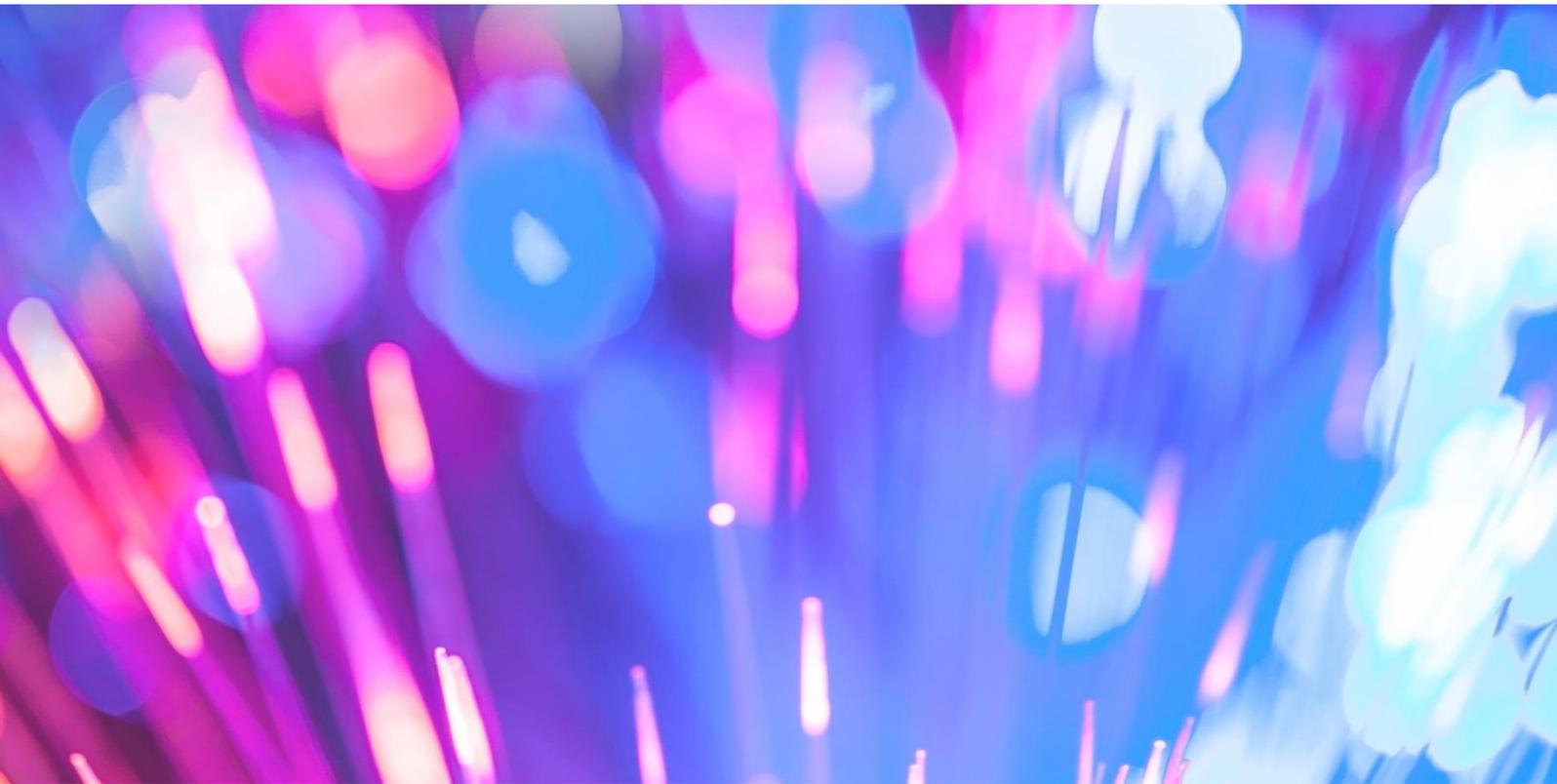# Searches: From Read to SNOMED CT

SNOMED CT in Primary Care project

Published September 2017

Version 1.0

Information and technology for better health and care

Contents

1. Introduction

The adoption of SNOMED CT in systems that are currently based on the Read codes will probably have greatest impact on functionality that incorporates a search (query).

In the GPSoC requirements, all supplier systems have to migrate existing searches (queries) to SNOMED CT. This document aims to provide information on how existing searches/queries may behave once migrated to SNOMED CT and what needs to be done to extend these to take account of the full UK Edition of SNOMED CT. For those with locally written searches in a separate local database solution, it provides insights into how such searches may be converted.

1.1 Audience

The intended audience is anyone who authors or updates searches in general practice systems or processes data from general practice that incorporates Read codes (this may be in a separate database). This document focuses on the principles underlying the actual change from Read to SNOMED CT and what this means for searches. Therefore end users, system suppliers, as well as teams with their own in house solutions, may also benefit from reading this document.

1.2 Context

Queries/searches/lists are various words used to describe the same thing: this document will use the words interchangeably. In essence, we are referring to the process of using Read codes to either select patients based on some criteria (e.g. all patients who have a diagnosis of Asthma), or identifying elements from the patients record (e.g. select all recorded blood pressure results in the last 3 months).

Such functionality includes:

- Reports. Reports are built on *searches* and are generally run through a specific reporting module. Suppliers and national organisations may provide the requirements for such reports.

- Mail merge letters, call/recall systems. *Searches* are used to 1. Identify the patients to include in the mail merge; and 2. specify data to extract from the record to include in the mail merge letter (e.g. observations in the record to send as part of the referral letter).

- Subsets / concepts / Formularies. In a number of systems it is possible to pre-define the clinical terms that are available for selection, often known as subsets or code lists. These may be used to restrict data entry in a template or in a report. In some systems these are referred to by name e.g. diabetic patients. These may be 'cherry picked' lists or may be created using a particular *search criteria*.

- Business protocols. In some systems it is possible to define a business rule that can be triggered either at a specific time or in response to particular data entry activity or sequence of events. These often use Read codes within a *search criteria*.

- Clinical Decision support tools such as CHA2DS2 -VASc, QRisk2

- Templates. *Searches* may be in use to identify the possible options for particular data items in data entry templates. They may use pre-defined code lists which may behave slightly differently when automatically converted from Read to SNOMED CT – for example a simple code list may become larger or smaller. Templates are often used as part of data quality to restrict the entries available for particular data items; it may be undesirable for this to include additional SNOMED CT concepts on conversion.

All the scenarios listed above share the common theme that they are currently driven by searches using Read codes and need to migrate to being driven by SNOMED CT concepts.

1.3 Scope

The document addresses the principles behind writing a simple search in SNOMED CT then looks how that query may be converted from Read to SNOMED CT and the consequences that may arise from that. Out of scope is the maintenance of queries specified in SNOMED CT following a new release of the terminology.

This is not a technical document and does not provide guidance on how to write searches either in a system or via SQL; it will provide an understanding of the way SNOMED CT and Read are different and how this plays out in the creation or migration of searches to SNOMED CT.

Separate technical documentation will be available on how a database currently used for data analysis, including data recorded using Read codes, may be updated to take account of the SNOMED CT terminology. This documentation will shortly be released on the NHS Digital website.

1.4 Further information

This document is provided by the SNOMED CT in Primary Care project within NHS Digital. Any questions arising from the content of this document or requests for further information can be sent to snomedprimarycare@nhs.net. Further resources, including a variety of recorded and live presentations, documents and short animations are available on our website. Readers may also be interested in the on-line look-up database providing the mappings from Read codes to SNOMED CT.

The NHS Terminology Service within NHS Digital also provide various resources to support those wishing to manage the transition of searches to take account of the more extensive and richer clinical terms provided by SNOMED CT. Of particular interest is:

- The NHS Data Migration pack of resources: this provides the mapping tables mentioned within this document.

- The UKTC Data Migration Workbench and the associated videos on its use: This is an Access database that can be used to map Read code lists to SNOMED CT, and help to extend that list to include all relevant SNOMED CT content. It can also be used to develop SNOMED CT code lists for particular search criteria.

- The SNOMED CT Request portal: requests for new terms and concepts in SNOMED CT can be made via this portal.

2. Fundamentals

Note. It is useful to understand some key principles of searches in SNOMED CT before continuing with the rest of the document.

2.1 Expressing searches - Expression Constraint Language

The international standard of SNOMED CT includes an expression constraint language (ECL) for use in writing queries. This is not a computer query language but a simple way of expressing sets of concepts that are of interest i.e. a 'syntax' with operators.

QOF 2018/19 will be defined in the Expression Constraint Language syntax

For example: To identify Asthma codes the syntax **<<Asthma (disorder)** will be used rather than using the current approach of H33..%.

Figure 1 shows the operators in the ECL which would be used to identify concepts to be included in a search.

Figure 1: Expression Constraint language (ECL)

ECL statement	Meaning
conceptid	this concept to be included in a query
<conceptid	the descendants of this concept to be included in a query but not the concept itself
<<conceptid	this concept and all its descendants to be included in the query
Then any of the above preceded by	
MINUS	exclude concept(s) in a query
OR	include concept(s) OR concept(s) in a query
AND	include concept(s) AND concept(s) in a query
^refsetId	Include all members of refset (the mechanism in SNOMED CT to identify a predefined subset of codes) in the query

The concepts themselves will be expressed using the following syntax:

ID | description |

This approaches allows the concept to be computer readable using the Concept ID and human readable using the description. The description in pipes | | can either be the Fully Specified Name (FSN) or the preferred term. NHS Digital generally uses the FSN.

Example: 233838001 | Acute posterior myocardial infarction (disorder) |

This syntax will be used throughout the rest of this document.

2.2 Mapping tables: maps from Read codes to SNOMED CT

NHS Digital publishes mapping tables that provide a SNOMED CT Concept ID plus Description ID for every Read code and term in the UK release of Read codes (separate tables for Read v2 and CTV3). They are designed to work in one direction only, from the legacy terminology to SNOMED CT. The maps have been used by system suppliers to map Read codes that are already in patient records to SNOMED CT concepts. This means that queries using SNOMED CT will still retrieve patient data that was entered prior to the move to SNOMED CT since each Read code in the record will have a SNOMED CT equivalent.

The maps have been created at description level i.e.:

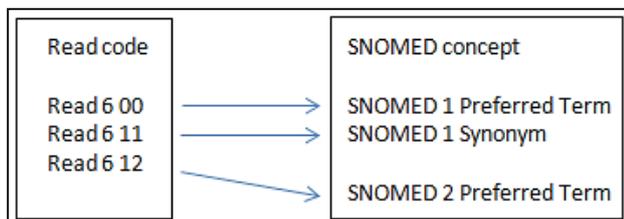**Read Code + Read Term code=SNOMED CT Concept + SNOMED CT Description**

Maps are at Read term level not Read code level to allow for the fact that within Read not all synonyms are true synonyms. Each Read synonym has been mapped to the best match SNOMED CT description, based on all the information in Read.

The maps can exhibit the following characteristics:

2.2.1 Non synonymous synonyms in Read (one-to-many)

Different synonyms of one Read code may map to different SNOMED CT concepts as illustrated in figure 2. This is due to synonyms not always being true synonyms in Read; and means that one Read *code* may map to more than one SNOMED CT concept

Figure 2: One Read code mapping to several SNOMED CT concepts

For example, the table below shows how H330 in Read v2 maps to SNOMED CT. This highlights that one Read code can map to several SNOMED CT concepts, in this example to three SNOMED CT concepts. This is due to the synonyms of Read not being true synonyms.

Figure 3: Maps from Read v2 to SNOMED CT - synonyms

Read code	Term code	Read term	SNOMED CT Concept id	SNOMED CT Description Id	SNOMED CT description
H330.	00	Extrinsic (atopic) asthma	389145006	1493645010	Allergic atopic asthma
	11	Allergic asthma	389145006	1483199016	Allergic asthma
	12	Childhood asthma	233678006	350147016	Childhood asthma
	13	Hay fever with asthma	233683003	350153016	Hay fever with asthma
	14	Pollen asthma	233683003	350154010	Pollen asthma

2.2.2 Duplicates (many-to-one)

Several Read codes may map to the same SNOMED CT concept as shown in Figure 4. This is usually due to deliberately introduced duplication in Read v2.

Figure 4: Maps from Read to SNOMED CT - duplicates
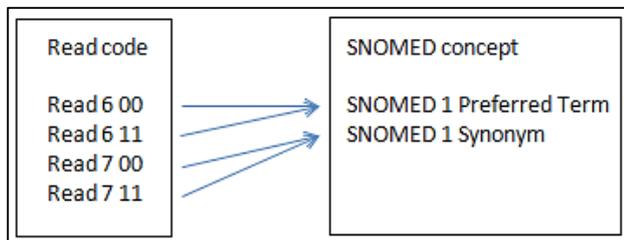
An example of duplicate Read Codes is H33.. 00 Asthma, H33z. 00 Asthma unspecified and H33zz 00 Asthma NOS which all clinically mean Asthma. These are all mapped to the SNOMED CT Concept ID: 195967001.

A further example is where a clinical term is found in more than one Read chapter (which was not possible in Read v2).For example, the table in Figure 5 below shows two Read codes which describe the same clinical concept but are in different chapters; these map to the same SNOMED CT concept.

**Figure 5: Maps from Read v2 to SNOMED CT - duplicates**

Read code	Term code	Read term	Conceptid	DescriptionId	Term
F004.	00	Meningitis – tuberculous maps	58437007	97107015	Tuberculous meningitis
A130.	00	Tuberculous meningitis maps	58437007	97107015	Tuberculous meningitis

2.2.3 Map to a higher level concept in SNOMED CT (many-to-one)

This scenario generally exists where there is no corresponding term (exact match) in SNOMED CT for the Read term. This is mainly because the original terms were introduced in the 1980s or 90s and are now agreed not appropriate to be included in a patient record and therefore are not found in SNOMED CT. In this case the Read codes map to SNOMED CT concepts that are more generic i.e. higher up the SNOMED CT hierarchy, as shown in Figure 6. This is often the case for terms containing NOS and NEC which are not present in SNOMED CT.

There are also some areas of Read that were created from ICD which have lots of codes related to things that it was felt important to count at some point in time: for example assault with a weapon, assault with a knife, assault with … In a clinical record for patient care we tend to just report assaulted. In this case, all these variations have been mapped to a single SNOMED CT concept as the large number of different types of assault is not replicated in SNOMED CT.

Figure 6: Read terms mapping to SNOMED CT concepts higher up the hierarchy
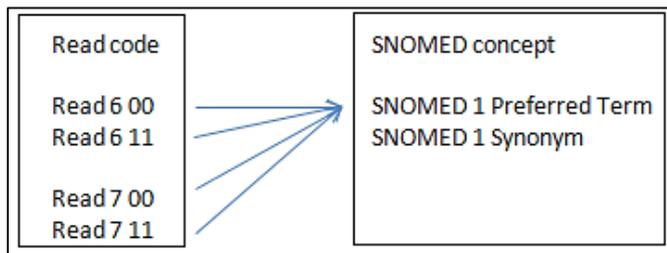
In the examples below in Figure 7, all Read codes map to SNOMED CT concepts that are higher up the hierarchy because the exact terms do not exist in SNOMED CT due to the terms not being appropriate for a modern terminology and are legacy terms from Classifications such as ICD-9 and ICD-10.

Figure 7: Mapping to SNOMED CT concepts higher up the hierarchy[1]

Read code	Term code	Read term	SNOMED CT Concept Id	SNOMED CT Description Id	SNOMED CT Description
F52z.	00	Otitis media NOS	65363002	108597015	Otitis media
Q4z..	00	Stillbirth NEC	237364002	355748018	Stillbirth
Eu32z	13	[X]Prolonged single episode of reactive depression	87414006	2951871015	Reactive depression
U102.	00	[X]Fall involving ice-skates, skis, roller-skates or skateboards	1912002	4301012	Fall

2.3 Maps to support running Read searches after transition

Once practices have moved to SNOMED CT, this will be the terminology that is entered into the patient record. To aid migration, NHS Digital have released a GP reference set (often called the GP subset) to be used when SNOMED CT is entered into a patient record. This enables systems to also add a Read code equivalent if one exists; the Read code will automatically be entered into the patient record. This is often referred to as dual coding. This approach allows a period of time for searches to be reviewed and extended to SNOMED CT since searches already written in Read will still run once the system is using SNOMED CT for data entry.

Figure 8 illustrates how these tables support dual coding in a patient record.

---

[1] Note in both of these examples, the other term codes for the Read codes map to different SNOMED CT concepts as per the synonym example but are not shown here for simplicity.

Figure 8: Use cases for the mapping files
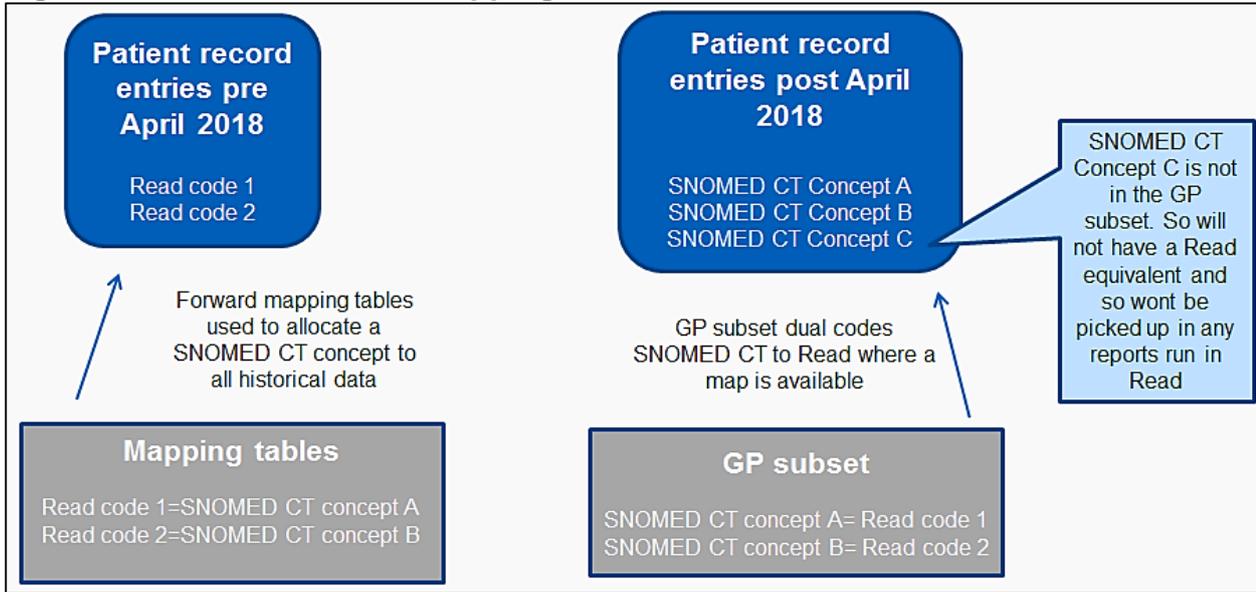
Note that when a SNOMED CT concept is entered into a patient record where there is not a Read equivalent, (illustrated by Concept C in the above example), then any searches run in Read will not find this data; it is important to keep this in mind when deciding priorities for updating searches to SNOMED CT.

The next section looks at writing a new search in SNOMED CT. Recall that due to the maps above, patient data originally recorded in Read will now have a SNOMED CT equivalent so a SNOMED CT search can still run over historical data. Also bear in mind that the principles and structure of the Read hierarchy is different to that in SNOMED CT.

3. Writing a new search in SNOMED CT

This section will focus on the principles behind writing a new search in SNOMED CT rather than migrating an existing Read defined search.

Searches in SNOMED CT are always undertaken within the system using the ConceptId only, as synonyms are true synonyms of the concepts.
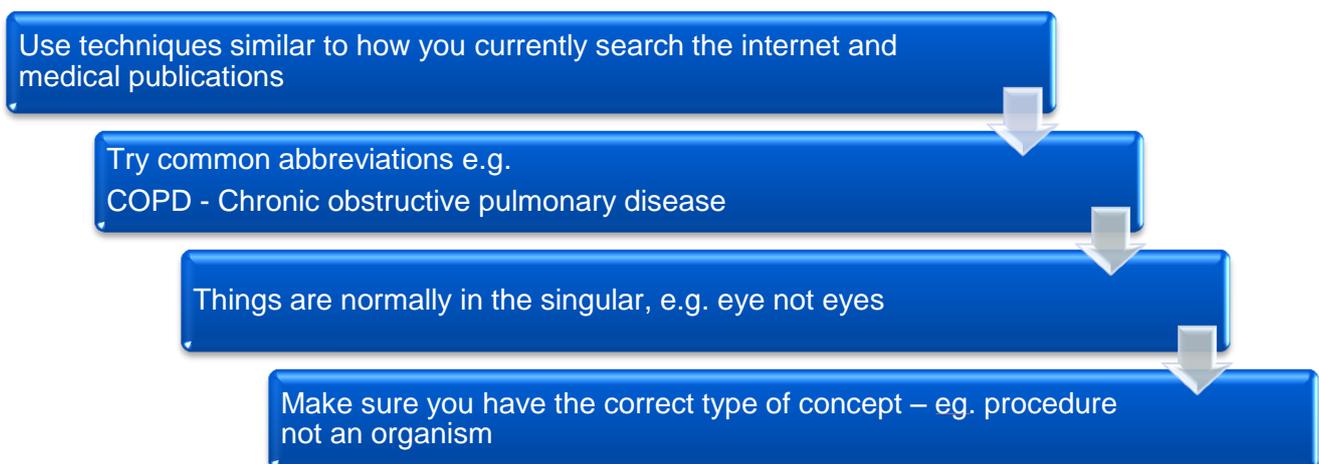
3.1 Finding concepts in SNOMED CT

The first stage in writing any search is to identify the concepts to include in that search.

Note. You may wish to quickly refresh your memory on the operators within the ECL in Figure 1 which indicates the approaches available for specifying a search.

SNOMED CT IDs do not hold any meaning, therefore searches cannot be based on the first digits of the code as may have been the case in Read searches; for example looking for all codes beginning with H33 to find all asthma codes. Instead, supplier interfaces will allow users to search for terms in a similar way to how they would search the internet, by simply typing what they want to find. Full word typing is often not required so typing the first few characters of each word may suffice and including words like the, and, of are often unnecessary to return the desired result. The word order is also often irrelevant

Figure 9: Tips for searching in SNOMED CT

Use techniques similar to how you currently search the internet and medical publications

Try common abbreviations e.g.
COPD - Chronic obstructive pulmonary disease

Things are normally in the singular, e.g. eye not eyes

Make sure you have the correct type of concept – eg. procedure not an organism

3.2 Using the hierarchies

In a search we need to ensure we have selected the appropriate concepts from the appropriate hierarchy. In SNOMED CT the hierarchies are used to group all the same types of codes together e.g. all procedures. It is important when selecting a concept to include in a search that you have not only checked the text, but have also ensured it is the right type of concept i.e. from the correct hierarchy.

Supplier interfaces may allow users to restrict searches to certain hierarchies for example just disorders. Recall there are 19 hierarchies in SNOMED CT as shown in Figure 10. The hierarchies that content most relevant for patient records will be under are:

- – Clinical finding
- – Procedure
- – Observable

– Stages and scales

– Situation with explicit context

The **clinical finding** concepts - represent result of a clinical observation (descriptive rather than with a value), assessment or judgment and include concepts used for diagnosis
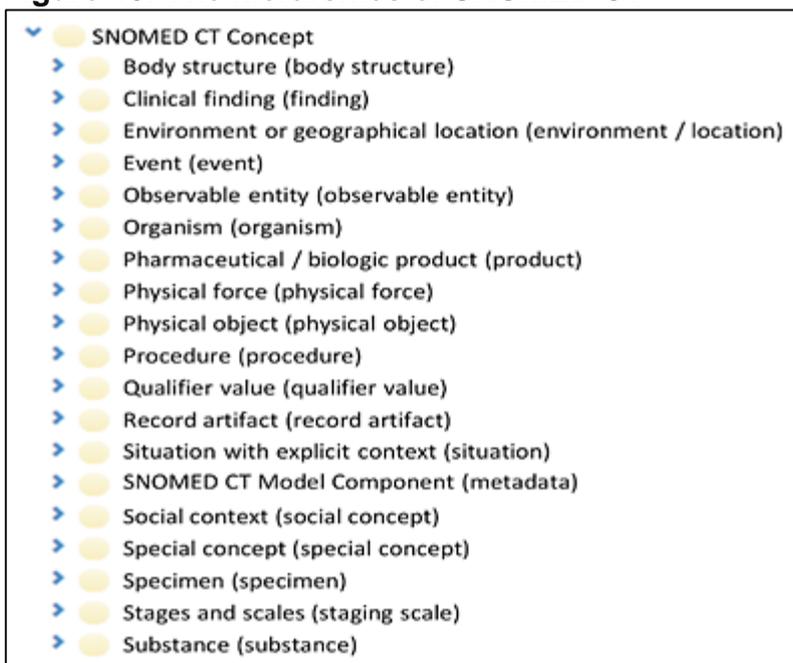
The **procedure** concepts - represent activities performed in the provision of health care. Includes invasive procedures, administration of medicines, imaging, education, therapies and administrative procedures

The **observable** concepts - Represents a question or assessment which can produce an answer or result (weight, blood pressure)

The **staging and scales** concepts – Represent Assessment scales and tumour staging systems

The **situation** concepts - represents concepts in which the clinical context is specified as part of the definition of the concept itself. For example, when the conditions are absent or in the past or relate to someone other than the subject of the record.

Figure 10: The hierarchies of SNOMED CT


```
⌄  ◯ SNOMED CT Concept
   ❯ ◯ Body structure (body structure)
   ❯ ◯ Clinical finding (finding)
   ❯ ◯ Environment or geographical location (environment / location)
   ❯ ◯ Event (event)
   ❯ ◯ Observable entity (observable entity)
   ❯ ◯ Organism (organism)
   ❯ ◯ Pharmaceutical / biologic product (product)
   ❯ ◯ Physical force (physical force)
   ❯ ◯ Physical object (physical object)
   ❯ ◯ Procedure (procedure)
   ❯ ◯ Qualifier value (qualifier value)
   ❯ ◯ Record artifact (record artifact)
   ❯ ◯ Situation with explicit context (situation)
   ❯ ◯ SNOMED CT Model Component (metadata)
   ❯ ◯ Social context (social concept)
   ❯ ◯ Special concept (special concept)
   ❯ ◯ Specimen (specimen)
   ❯ ◯ Stages and scales (staging scale)
   ❯ ◯ Substance (substance)
```

3.3 Using the |is a| relationships

Concepts within the same hierarchy are linked together (where appropriate) as parent and child concepts via a relationship called the |is a| relationship. Each concept must have at least one parent concept. It is the |is a| relationships that form the basis of our searches.

Figure 11 shows some of the |is-a| relationships for bacterial pneumonia. This illustrates that SNOMED CT is more of a network of links than a linear structure. It also demonstrates why it is more important to search for a particular concept rather than trying to find a concept by navigating through the hierarchy from the top.

Figure 11: Bacterial pneumonia – going up each parent via the |is a| relationships
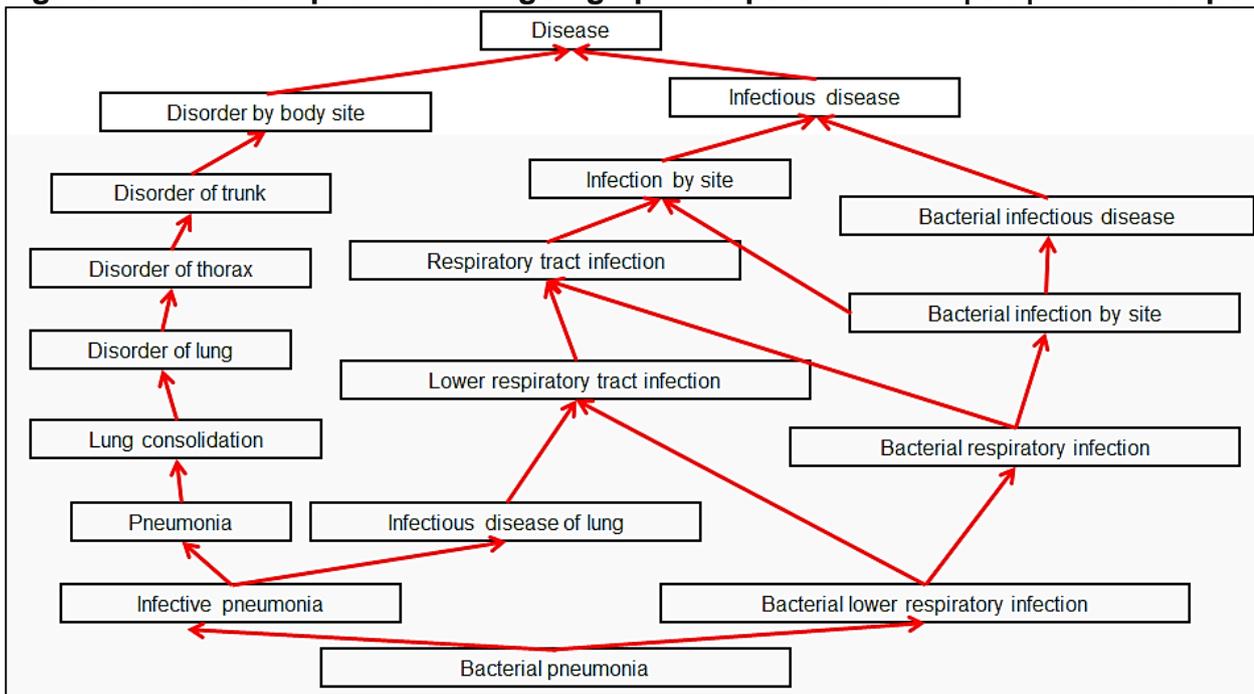
The relationship structure of SNOMED CT can be used to create searches for a concept and all descendants that meet our requirements; note that all descendants of a concept are true subtypes of it. For example SNOMED CT would not place Not Constipated as a subtype of Constipated as is the case in Read v2. We will come back to this later in this document.

Example

Creating a search for patients who have had any type of appendectomy. Searching for Appendectomy reveals there is a concept in SNOMED CT of 80146002 |Appendectomy (procedure)|. This is in the procedure hierarchy (so correct type) and has eight children, some of which also have children (as denoted by < before the concept name in the list of children in the search results. This changes to ⌄ when clicked on to reveal the children). See Figure 12 where the children have all been expanded to show full set of concepts linked to 80146002 |Appendectomy (procedure)| via the |is-a| relationships in SNOMED CT.

Note that due to the structure of SNOMED CT, concepts can have any number of parents and therefore a concept can appear in more than one place in a hierarchy tree view. . There is one such example in Figure 12 highlighted in blue. If the search required that 174041007 |Laparoscopic emergency appendectomy (procedure)| should not be included and in the supplier interface this was de-selected, then both instances should be unselected.

To indicate all concepts that are descendants of 80146002 |Appendectomy (procedure)| plus the concept itself, then this should be expressed in ECL as:

<<80146002 |Appendectomy (procedure)|

Figure 12: Appendectomy and descendants
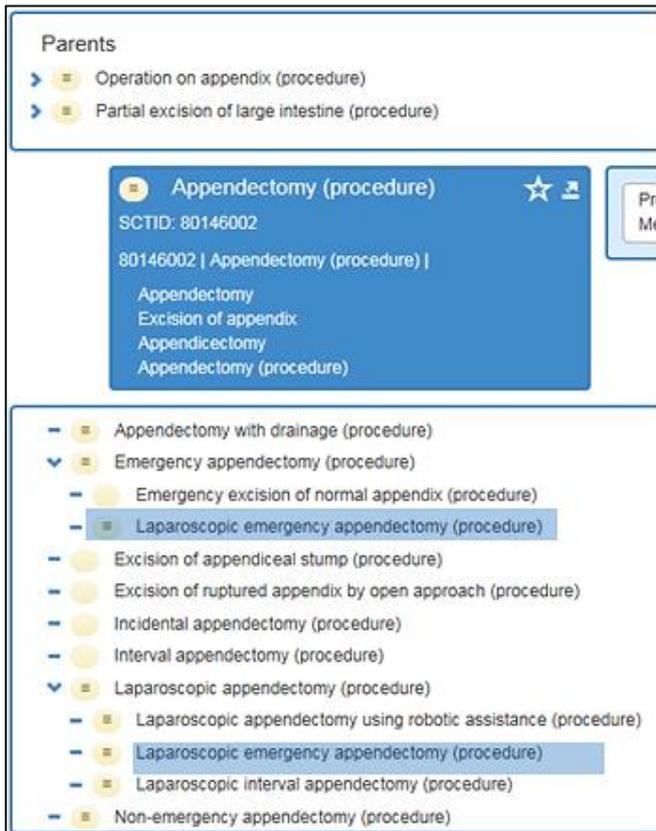
This shows a very simple way to start to structure searches; users are encouraged to become familiar with the hierarchies and to explore concepts that are common within their work setting. System suppliers will be providing training and documentation on their interfaces and how to structure searches in their environments.

Note searches can be more complex by incorporating a number of concepts and all their children with the use of AND and OR (+), or excluding content using MINUS.

Example:

421720008 |Spray dose form| OR 7946007 |Drug suspension|

This represents a drug dose form that is either a spray or a suspension or both

4. Differences from Read

The principles behind how the hierarchies are formed are different in SNOMED CT to those in Read. This section highlights the main differences.

4.1 Hierarchy structure

Each concept within a hierarchy in SNOMED CT is a true subtype of its parent concept. In Read this is not always the case, for example in Read v2, 19C2. Constipated and 19C1. Not Constipated are both subtypes of 19C.. Constipation. In SNOMED CT these two concepts still exist but are not under the same parent concept. In SNOMED CT the parents of the two concepts are:

SNOMED CT concept	Parents
14760008 |Constipation (disorder)| Syn: Constipated	128524007 |Disorder of colon (disorder)| 81120009 |Functional disorder of intestine (disorder)|
162081000 |Not constipated (finding)|	300380005 |Finding of frequency of defecation (finding)|

Both are still in the clinical findings hierarchy but one is a disorder and the other is a statement we can make about the patient and is a finding. In a Read v2 report it is highly likely Not constipated has been excluded from a search for patients with constipation. In SNOMED CT 162081000 |Not constipated (finding)| would not need to be removed from a search looking for patients who have constipation as it is not a subtype of it. In fact, 14760008 |Constipation (disorder)| has 15 children, some of which are shown in Figure 13 illustrating that all are a subtype of constipation.

Figure 13: Some of the child concepts of 14760008 |Constipation (disorder)|

Children (15)
- Acute constipation (disorder)
- Chronic constipation (disorder)
- Constipation - functional (disorder)
- Constipation alternates with diarrhea (disorder)
- Constipation by outlet obstruction (disorder)
- Constipation due to atony of colon (disorder)
- Constipation due to neurogenic bowel (disorder)
- Constipation due to spasm of colon (disorder)
- Drug-induced constipation (disorder)
- Irritable bowel syndrome characterized by constipation (disorder)
- Irritable bowel syndrome variant of childhood with constipation (disorder)

4.1.1 Semantic tag

Each Fully Specified Name (FSN) term ends with a "semantic tag" in parentheses. The semantic tag indicates the semantic category to which the concept belongs (e.g. clinical

finding, disorder, procedure, organism, person, etc.). The "semantic tag" helps to distinguish one concept from another especially where descriptions are very similar. For example: 129003000 |Dressing (observable entity)| meaning someone is able to dress themselves and 3895009 |Application of dressing (procedure)| meaning to dress a wound.

Figure 14 shows the semantic tags and sub hierarchy tags. For example, the top level hierarchy 'clinical finding' hierarchy has two sub hierarchy tags of 'finding' and 'disorder'. It is these tags that appear in brackets after the Fully Specified Name. In some cases there may be a second level as shown in the third column.

Figure 14: Semantic tags and sub hierarchy tags

Top level hierarchy	Sub hierarchy tag	Sub hierarchy 2 tag
body structure	(body structure)	(cell)
body structure	(morphologic abnormality)	
clinical finding	(disorder)	
clinical finding	(finding)	
environment or geographical location	(environment)	
environment or geographical location	(geographic location)	
Event	(event)	
linkage concept	(attribute)	
linkage concept	(link assertion)	
observable entity	(observable entity)	
Organism	(organism)	
pharmaceutical / biologic product	(product)	
physical force	(physical force)	
physical object	(physical object)	
Procedure	(procedure)	
Procedure	(regime/therapy)	
qualifier value	(qualifier value)	
record artifact	(record artifact)	
situation with explicit context	(situation)	
social context	(social concept)	
social context	(person)	
social context	(ethnic group)	
social context	(racial group)	
social context	(religion/philosophy)	
social context	(occupation)	
social context	(life style)	

social context	(family)	
special concept	(inactive concept)	
special concept	(namespace concept)	
special concept	(navigational concept)	
Specimen	(specimen)	
staging and scales	(assessment scale)	
staging and scales	(tumor staging)	
staging and scales	(staging scale)	
Substance	(substance)	

4.2 SNOMED CT is not ambiguous

SNOMED CT is designed to be precise and logical; this leads to a number of differences from Read. In the main, these will have been addressed by the mapping from Read to SNOMED CT; however when writing a new search one needs to be aware of these when searching for the concepts to include.

A full description of the types of differences between Read and SNOMED CT is provided in the Data Quality Guidance. A few are provided here to encourage readers to reference the Data Quality Guidance if they have not already done so or have difficulty in finding concepts for terms previously used in Read.

One should also check the content in pre-existing templates and how that will be mapped to SNOMED CT. Our early investigations indicate that sometimes incorrect Read codes have been used because a particular code might not have existed: for example *Under the care of a hospice* does not exist in Read and just Hospice was provided in the template. This must have been accounted for in any Read queries so will be accommodated in a converted SNOMED CT query; however it should be noted that more applicable codes may exist in SNOMED CT and data using the correct concept may come into the patient record from elsewhere. If such data quality issues are not resolved, this could result in some patients not being identified in new searches developed in SNOMED CT.

4.2.1 SNOMED CT does not allow concepts such as NOS, NEC 'other'

NOS, NEC, other

In Read, terms with NOS, NEC and' Other' are used as a 'catch all' group. In SNOMED CT no such groups exist as SNOMED CT aims to be precise and clear. To illustrate two examples below are shown:

F52z. 00 Otitis media NOS   maps to 65363002 | Otitis media |

Q4z.. 15 Stillbirth NEC   maps to 237364002 | Stillbirth |

Both Read codes map to a term higher up the hierarchy in SNOMED CT indicating that a non-specific concept can be recorded but will not have NOS, NEC or 'Other' in the name.

4.2.2 Different meanings in different settings

There are some commonly used codes in Read which may mean different things in different settings. Where it is not safe to use such codes for interoperability, these generally map but to inactive codes in SNOMED CT. For example both Read codes below map to an inactive SNOMED CT concept so will not be available for future data entry; this is because it is not clear what Clinic A is and this may be differ from practice to practice.

9OS3. 00 Clinic A monitoring default

5. Migrating Read searches to SNOMED CT

There are hundreds of searches/reports/queries in existence that are in Read. These will continue to function after the transition to SNOMED CT but only over data that can be dual coded.

All system suppliers will have mapped historical Read codes in patient data to SNOMED CT concepts so any query written in SNOMED CT will work over historical patient data.

Within NHS digital, work has been undertaken to understand how searches can be converted any observations from using the mapping tables. This section describes the understanding gleaned from this and how this can be used when reviewing or converting searches locally.

In the main, searches in supplier systems will be migrated to SNOMED CT. However, it is not clinically safe for such searches to be automatically extended to take account of SNOMED CT concepts that do not have equivalent codes in Read. National searches provided within the system by suppliers will be converted and extended.

There are two ways to approach migrating Read searches:

- Take the previous **query definition** and write a **new SNOMED CT query** from scratch

- A **Mapping approach**: Take the previous codes that are in *scope of the original query* and *map* each of these to SNOMED CT using the mapping tables. Then decide whether this needs extending to include all the concepts in SNOMED CT now or at a later date.

5.1 Query definition approach

By query definition we mean going back to what is it we are interested in reporting on. Let us not worry about the Read codes that were used in the past but start from scratch and look to see which SNOMED CT concepts we should include.

The approach requires the use of tools including a SNOMED CT browser; producing a list of the codes to search for. This approach will use the whole of SNOMED CT (i.e. not just SNOMED CT that has a map from Read) so is likely to include concepts that do not have a Read equivalent. However, this does not matter as these codes will not currently be in records, but could start to be as users start to use SNOMED CT or receive data from outside the practice. Long term, this is a more robust approach.

When using this approach, users should also understand how any historical data is mapped from Read to SNOMED CT to ensure everything applicable is captured by the new search. To understand any additional codes needed to capture historical Read content then the mapping tables can be used to also map the codes of the original Read query. Examination of any concepts not in the newly written query should be undertaken to understand why these do not appear and are they needed to ensure historical content is found.

5.2 Mapping approach

This requires use of the mapping tables. The original query needs to be expanded to provide a list of all the Read codes plus terms previously included; each is then mapped to the SNOMED CT equivalent using the mapping tables.

If the original Read query is expressed only as Read codes and all descendants, then each expression code needs to be expanded to include all the Codes in scope with all the terms for each code. Each Code with term then needs to be expressed as the SNOMED CT equivalent found using the appropriate mapping table.

Note. It may be tempting to take each term in a Read query expression and substitute it with the equivalent concept in SNOMED CT. For example:
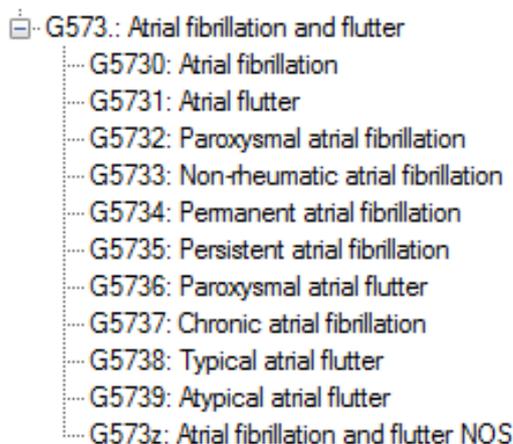
<div align="center">H33..% Asthma to << 195967001 | Asthma (disorder) |</div>

However, differences in the structure of the hierarchies in SNOMED CT to those in Read and the structure of the mapping tables highlighted above mean this may not work as expected.
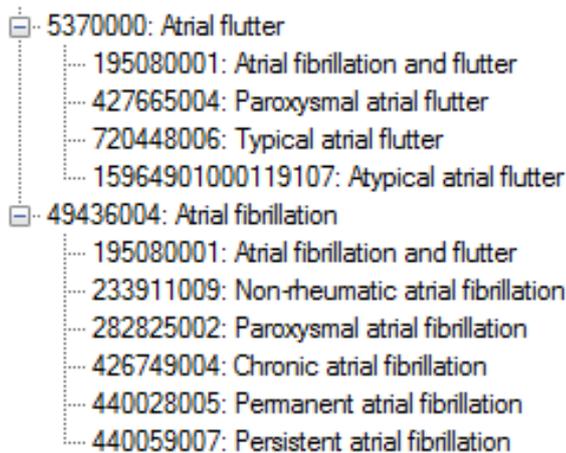
Example: Atrial fibrillation

In Read v2, the hierarchy structure has both "Atrial fibrillation" and "Atrial flutter" as children of "Atrial fibrillation and flutter" as displayed in Figure 15.

Figure 15: Read v2 hierarchy for Atrial fibrillation

```
⊟ G573.: Atrial fibrillation and flutter
    ├─ G5730: Atrial fibrillation
    ├─ G5731: Atrial flutter
    ├─ G5732: Paroxysmal atrial fibrillation
    ├─ G5733: Non-rheumatic atrial fibrillation
    ├─ G5734: Permanent atrial fibrillation
    ├─ G5735: Persistent atrial fibrillation
    ├─ G5736: Paroxysmal atrial flutter
    ├─ G5737: Chronic atrial fibrillation
    ├─ G5738: Typical atrial flutter
    ├─ G5739: Atypical atrial flutter
    └─ G573z: Atrial fibrillation and flutter NOS
```

Whereas in SNOMED CT "Atrial fibrillation and flutter" is a child of both "Atrial fibrillation" and "Atrial flutter" as displayed in Figure 16. If a query had been converted at parent level and just assumed all children were the same both "Atrial fibrillation" and "Atrial flutter" would not be included in a query converted in this way.

Figure 16: SNOMED CT hierarchies of Atrial fibrillation and Atrial flutter


```
⊟ 5370000: Atrial flutter
        195080001: Atrial fibrillation and flutter
        427665004: Paroxysmal atrial flutter
        720448006: Typical atrial flutter
        15964901000119107: Atypical atrial flutter
⊟ 49436004: Atrial fibrillation
        195080001: Atrial fibrillation and flutter
        233911009: Non-rheumatic atrial fibrillation
        282825002: Paroxysmal atrial fibrillation
        426749004: Chronic atrial fibrillation
        440028005: Permanent atrial fibrillation
        440059007: Persistent atrial fibrillation
```

Therefore, any queries written in Read need to be converted on a code by code basis. In the example above, all of the descendants of G573 would be mapped to SNOMED CT independently to find SNOMED CT equivalents.

5.3 Reflect on the effect from historical Read Data

In Section 2 we saw how the mapping to SNOMED CT addresses things like duplicates in the Read codes. So what if, for example, the original query had missed one of these duplicates? The following illustrates how the results of your new query may differ from the searches when previously run in Read, the reasons why and how to understand how any differences may occur.

5.3.1 Steps to using the mapping tables to identify any differences

One way to identify additional historical records with Read codes that may now be returned when the search is written in SNOMED CT is as follows:

Step 1 – expand to all Read codes referred to within the expression: code%

Step 2 – ensure you have the Read terms as well

Step 3 – use the mapping tables to find the SNOMED CT concept for each code with term

Step 4 – NOW use the mapping tables to identify any additional Read codes that would be included in the SNOMED CT query as they map to the same SNOMED CT concepts. These would appear as they **also** map to the SNOMED CT concepts in Step 3.

Example: Using the Read v2 QOF Diabetes mellitus as an example (version 36) in http://content.digital.nhs.uk/media/24557/Diabetesv360/pdf/Diabetes_v36.0.pdf[2]

Starting point – the QOF cluster as expressed in Read v2 is:

```
C10.. , C109J , C109K , C10C.
, C10D. , C10E.% , C10F.%
(excluding C10F8) , C10G.% ,
C10H.% , C10M.% , C10N.% ,
PKyP. , C10P.% , C10Q.
```

---

[2] Note mapping tables may have changed since this example was created so this exact output may not be replicated locally.

Step 1 - expand all the % code expressions. For example C10N% actually represents 3 Read codes

C10N.	00	Secondary diabetes mellitus
C10N0	00	Secondary diabetes mellitus without complication
C10N1	00	Cystic fibrosis related diabetes mellitus

Do this for all instances of % in your query expression; in this example this results in 73 Read codes in total.

Step 2 –express all Read codes as code and term code pairs. For example: C109J has three synonyms:

C109J00	Insulin treated Type 2 diabetes mellitus
C109J11	Insulin treated non-insulin dependent diabetes mellitus
C109J12	Insulin treated Type II diabetes mellitus

Do this for all 73 Read codes, this results in 153 Read code and term pairs.

Step 3 – Use the mapping tables to find the equivalent SNOMED CT concepts for each Read code and term pair.

Using the examples above, the tables in Figure 16 and 17 show which SNOMED CT concepts and descriptions these map to. In the first table we can see that of the three Read codes, two map to the same SNOMED CT concept; in the second table the Read codes map to just one SNOMED CT concept (though to different descriptions).

Figure 16: maps from C10N% to SNOMED CT

Read code	Term code	Read term	SNOMED CT Conceptid	SNOMED CT DescriptionId	Term
C10N.	00	Secondary diabetes mellitus	8801005	15518018	Secondary diabetes mellitus
C10N0	00	Secondary diabetes mellitus without complication	8801005	15518018	Secondary diabetes mellitus
C10N1	00	Cystic fibrosis related diabetes mellitus	426705001	2674608014	Diabetes mellitus associated with cystic fibrosis

Figure 17: maps from C109J to SNOMED CT

Read code	Term code	Read term	Conceptid	DescriptionId	Term
C109J	00	Insulin treated Type 2 diabetes mellitus	237599002	2967820013	Insulin treated type 2 diabetes mellitus
	11	Insulin treated non-insulin dependent diabetes mellitus	237599002	1223148019	Insulin treated non-insulin dependent diabetes mellitus
	12	Insulin treated Type II diabetes mellitus	237599002	1223147012	Insulin treated Type II diabetes mellitus

Note. Although this table shows all synonyms of the same Read code mapping to the same SNOMED CT concept this is NOT always the case.
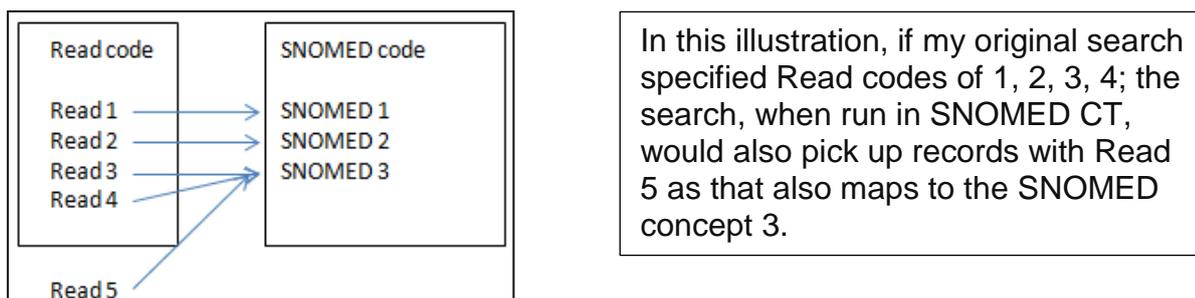
Using the mapping tables in this way for all 153 Read code and term pairs results in 67 distinct SNOMED CT concepts.

Note, however, that this translation does not use all of the SNOMED CT concepts that relate to diabetes so is of use only while dual coding continues for all data entered: it omits SNOMED CT codes that are clearly types of diabetes (e.g. Houssay's or Feigenbaum-Bergeron-Richardson syndromes, or Diabetes mellitus due to cystic fibrosis), but that will never be reached as the direct translation of any prior Read code since they did not exist in Read. This is illustrated later in the section on 'Extending the migrated search'.

Step 4 -use the SNOMED CT converted query and check if any other Read codes map to the list of SNOMED CT concepts in step 3; if they do any records containing these Read codes will also be in search results. (Illustrated in Figure 18).

As historical Read data will have SNOMED CT codes added to enable new SNOMED CT searches to run over such historical data, and we know that more than one Read code can map to the same SNOMED CT concept, we need to check if this is the case in our query. Generally this should be a more correct query, however the non-synonyms in Read have highlighted some interesting conversions as these were often not robustly checked in the original Read query, and often could not be excluded in system searches.

Figure 18: Using the mapping tables to migrate codes from Read to SNOMED CT

In this illustration, if my original search specified Read codes of 1, 2, 3, 4; the search, when run in SNOMED CT, would also pick up records with Read 5 as that also maps to the SNOMED concept 3.

Going back to the example of converting the QOF Diabetes from Read to SNOMED CT illustrates this.

The specification started with: 73 Read codes
This mapped to: 67 SNOMED CT concepts
The 67 SNOMED CT concepts are mapped to by: 138 Read codes
Note. This is an extreme example which results from historical data containing insulin dependent diabetes, and in practice may not result in a significant difference in patients returned as it is likely records will contain the old code AND the new code.

To help understand why there are these differences, let's consider Read code C10E9 which was part of the original specification. Figure 19 below shows that all synonyms of this Read code map to the same SNOMED CT concept 190372001. To identify if any other Read codes map to this concept, look to see which Read codes map to concept 190372001. Figure 20 shows that C1089 also maps to SNOMED CT 190372001. This means all patients with C1089 in their records would also be found by this converted query, which was not the case in the original Read search.

Figure 19: Mapping Read code C10E9 to SNOMED CT

Read code	Term code	Read term	Conceptid	DescriptionId	Term
C10E9	00	Type 1 diabetes mellitus maturity onset	190372001	292553010	Type 1 diabetes mellitus maturity onset
	11	Type I diabetes mellitus maturity onset	190372001	292551012	Type I diabetes mellitus maturity onset
	12	Insulin dependent diabetes maturity onset	190372001	292552017	Insulin-dependent diabetes maturity onset

Figure 20: using the SNOMED CT concept 190372001 to establish if additional Read codes will be identified in historical data

Read code	Term code	Read term	Conceptid	DescriptionId	Term
C1089	00	Insulin dependent diabetes maturity onset	190372001	292552017	Insulin-dependent diabetes maturity onset
	11	Type I diabetes mellitus maturity onset	190372001	292551012	Type I diabetes mellitus maturity onset
	12	Type 1 diabetes mellitus maturity onset	190372001	292553010	Type 1 diabetes mellitus maturity onset
C10E9	00	Type 1 diabetes mellitus maturity onset	190372001	292553010	Type 1 diabetes mellitus maturity onset
	11	Type I diabetes mellitus maturity onset	190372001	292551012	Type I diabetes mellitus maturity onset
	12	Insulin dependent diabetes maturity onset	190372001	292552017	Insulin-dependent diabetes maturity onset

Note: as the maps from Read to SNOMED CT are from Read code and term code to SNOMED CT concept and description, it can be the case that not all the terms of a Read code map to the particular SNOMED CT concept. Hence not every record with a specific Read code may be returned in a search, if it has a different term Id. This is in the example illustrated below.

If we also consider the concept 44054006 |Diabetes mellitus type 2 (disorder)| both the following Read codes map to that concept and are within the original specification:

C10F. 00   Type 2 diabetes mellitus
C10F. 11   Type II diabetes mellitus

Figure 21 shows other Read codes, not part of the original specification, which also map to 44054006 |Diabetes mellitus type 2 (disorder)| as they have the same meaning clinically. This illustrates that not all synonyms of Read code C109. map to 44054006 |Diabetes mellitus type 2 (disorder)|, and the Read code C1001 also maps to this concept. The preferred term C1001 00 "Diab.mell.no comp. – adult" maps to 111552007 |Diabetes mellitus without complication (disorder)| which is not a concept in our converted query specification.

Figure 21: all Read codes and terms that map to 44054006 |Diabetes mellitus type 2 (disorder)|

Read code	Term code	Read term	SNOMED CT Conceptid	SNOMED CT DescriptionId	SNOMED CT Description
C1001	11	Maturity onset diabetes	44054006	493774016	Type II diabetes mellitus
C1001	12	Non-insulin dependent diabetes mellitus	44054006	493773010	NIDDM - Non-insulin dependent diabetes mellitus
C109.	00	Non-insulin dependent diabetes mellitus	44054006	493773010	NIDDM - Non-insulin dependent diabetes mellitus
C109.	11	NIDDM - Non-insulin dependent diabetes mellitus	44054006	493773010	NIDDM - Non-insulin dependent diabetes mellitus
C109.	12	Type 2 diabetes mellitus	44054006	197761014	Type 2 diabetes mellitus
C109.	13	Type II diabetes mellitus	44054006	493774016	Type II diabetes mellitus

The key point being illustrated here is that when a search is migrated to SNOMED CT on a term by term basis then the search may return historical records with Read codes that were not in the original Read specification, because those Read codes map to the same SNOMED CT concepts.

This illustrates reasons why the search results may differ, and will help anyone responsible for writing searches to be able to understand and identify why they may get a different result for a particular search after the transition to SNOMED CT.

As mentioned previously, identifying records with so many extra Read codes, as in the example above, is extreme and was mostly due to QOF only accepting certain codes for payment. However, the converted query is clinically more correct. In most cases where this phenomenon occurs, the 'extra' data being returned is clinically appropriate to the query, suggesting that the original query specification may have been incomplete or intentionally sub-optimal for say data quality reasons.

5.4 Extending a migrated search

In the example above, the original Read query was migrated to SNOMED CT by translating each Read code to SNOMED CT via the mapping tables. Therefore, the new SNOMED CT query is restricted to SNOMED CT concepts that have a Read equivalent. SNOMED CT is richer and more extensive than Read, and so inevitably there will be concepts from SNOMED CT that should be in the search specification that aren't.

Let's look at the concept 73211009 |Diabetes mellitus (disorder)| in a SNOMED CT browser; Figure 22 shows there are16 direct children and many of these also have children as denoted by >. In fact there are 125 active descendants of 73211009 |Diabetes mellitus (disorder)|; the converted query found 67.

Figure 22: 73211009 |Diabetes mellitus (disorder)| as seen in a browser

Depending on what the search is intended for, it is highly likely that most if not all of these additional concepts will be required in a search for diabetes. For example 75682002 |Diabetes mellitus caused by insulin receptor antibodies (disorder)| is a descendant of 73211009 |Diabetes mellitus (disorder)| but as it does not have a Read equivalent it would not be in query converted from the original Read codes.

Our recommendation is therefore when reviewing a search, to re-visit the original specification and write the query in SNOMED CT, then inspect the original Read search to see if this identifies other SNOMED CT concepts required in order to retrieve historical data. This may be because previously less than ideal Read codes have been used due to the codes not being in Read, or as in our example here specific Read codes have been not been included previously.

Some supplier solutions may suggest as part of their search writing functionality additional codes that may be appropriate to include; however this is difficult to do automatically for complex searches and as the SNOMED CT hierarchy structure differs to that in Read any such suggestions should always be reviewed.

5.4 Impact of migrated searches

Moving from Read to SNOMED CT involves changes to historical patient records as well as queries, templates etc. Future content entered into patient records may have been selected having the SNOMED CT hierarchy in view and hence new content available for selection. This all contributes to possible differences in the results of migrated searches. As shown in this document, it is therefore expected that search results from before and after will differ.

6. Some General Points

Remember, searches in SNOMED CT are always undertaken using the Concept Id and hence will identify data containing any of the descriptions of the concept.

6.1 Reviewing existing searches after the transition

It is advised that where possible, practices start to review their local searches and identify those that are used frequently; these should then be given priority for review after transition and extending them to take account of the additional content in SNOMED CT if appropriate.

If in-house database solutions are used and there are searches within these that are outside the principle supplier systems, then these need to be converted to be expressed using SNOMED CT. Due to dual coding, there is time to make such changes depending on which clinical areas new SNOMED CT content may start to be used. Documentation aimed at the technical teams who manage these will shortly be available on the NHS Digital Delen website and address building the terminology tables for use within a local data warehouse.

6.2 Searches housekeeping

An updated version of SNOMED CT is released in the UK twice a year on 1st April and 1st October; this should be incorporated in supplier systems within two months of release. Each SNOMED CT release may have new content added (as with Read) or existing content can be retired (known as being made inactive – see next section). Any historical data containing inactive content will still be identified in searches in supplier systems. Those responsible for searches therefore need to be mindful that searches need to be periodically checked to ensure they still include all relevant content.

6.2 Inactive SNOMED CT Concepts

One of the shortfalls of Read v2 was that Read codes that were added years ago could not be amended when they were known to be no longer correct. As clinical understanding changes or terms become outdated this has led to inappropriate content being present in Read v2 'for ever'. One of the benefits of SNOMED CT is that such content can be retired. However, those who write searches have not had to deal with such situations before.

One approach may be that locally there is a process where such codes are batch replaced in historical records (with clinical oversight). Note that all systems provide audit trails and so if batch processes are undertaken it is possible to retrieve the original codes.

Another approach is to leave such content in the records and suppliers have a mechanism to retrieve such inactive content. Within your system search functionality you can decide whether or not to retrieve inactive content in a search. It is advised that, unless there is a known specific reason not to, that users always include inactive content in searches. Supplier systems should also allow users to identify and specify inactive content in the searches written. Again, it is likely that this can be switched on or off.