INTERNATIONAL HEALTH TERMINOLOGY
STANDARDS DEVELOPMENT ORGANISATION
UK Terminology Centre (UKTC)

NHS
Connecting for Health

THE CLINICAL TERMS VERSION 3
(THE READ CODES)

MANAGING CHANGE:
DESCRIPTION CHANGE FILE

APRIL 2008

Purpose of this document

This document is one of a series that, taken together, describe the contents, structure and function of Clinical Terms Version 3 (The Read Codes).

This introduction is intended to provide information on Clinical Terms Version 3. It is also a guide to the other available documents each of which is updated independently. For this reason, different chapters may have different version numbers.

INFORMATION

Distribution	On request
Author	UK Terminology Centre
Further Copies From	UK Terminology Centre Helpdesk Service Support Unit Prospect House, Floor 2 Fishing Line Road Redditch Worcestershire B97 6EW Tel: +44 (0) 1392 206248 Fax: +44 (0) 1392 206945 E-mail: datastandards@nhs.net Internet: http://www.connectingforhealth.nhs.uk/systemsandservices/data/readcodes
Date of Issue	April 2008
Reference Number	173v1.0

Published by the UK Terminology Centre

Table of Contents

1. Management Summary

Clinical Terms Version 3 (The Read Codes), or CTV3, is a very large clinical terminology which can be used to describe patient care in computerised clinical record systems. Its content changes at each release, mostly with additional terms and codes to meet newly discovered needs. However, some of the established content of CTV3 needs to be **corrected** at each release because of:

- New clinical knowledge
- Limitations in the coding scheme of earlier versions of Clinical terms (e.g. Read Codes Version 2 and the 4-byte set). All these terms and Read Codes are part of the Clinical Terms Version 3. In these schemes, terms with different meanings often had to share the same Read Code, a feature which goes against the principles of CTV3.
- Errors discovered in the new content of CTV3 which occur because of the size of the product (and despite the most rigorous quality assessment available in the world of terminology development!).

Other terminologies and classifications have either changed only rarely (e.g. once a decade), in which case the scheme becomes increasingly irrelevant to current practice as time passes by, or have ignored the effect that change has on users. This may result in searches and reports either failing to pick up all patients who should fit the query, or including others who do not.

This document 'Managing Change in Clinical Terms Version 3 (Read Codes) – Description Change File' describes an approach to managing change as a shared responsibility between terminology providers such as the UK Terminology Centre, system developers and terminology users (typically clinicians). It describes how a new release file, the ***'Description Change File'*** **(referred to hereafter as the 'DCF')** can be used to improve information retrieval for clinical records. **Information retrieval** includes querying a patient database for decision support on individual patients, and population queries for clinical and management audit.

The DCF works by suggesting a new Read Code for each Clinical term which has been found to be wrongly associated with its previous Read Code. For example, if *'Cardiac rupture'* was found to have the same Read Code as '*Myocardial infarct'*, then the relationship between *'Cardiac rupture'* and that Read Code would be deleted from CTV3. *'Cardiac rupture'* would be then be associated with a new Read Code in CTV3. Lastly, an entry is made in the DCF suggesting that the user replaces the old Read with the new Read Code.

To be sure that the clinical record remains **clinically correct and medico-legally useful**, these changes would be made according to a strict protocol.

The first safety principle is to distinguish the ***selected*** Read Code from the Read Code ***analysed*** in the clinical record. When first chosen, these are identical. The record of the user's ***selected*** code should always be kept unchanged, along with the user's selected term and term identifier. Retaining the selected Read Code is important both to maintain the medico-legal

status of the electronic record and also to act as an anchor for future changes suggested by the UK Terminology Centre. However, the Read Code used for analysis may be changed (updated) in order to retrieve records correctly (see Diagram 1).

The second safety principle is to ensure that any suggested changes are adequately considered before being made. This is made tractable by distinguishing four different classes of change. These are changes due to:

- **Redundancy**. When a concept is found to have two or more Read Codes. The terms are removed from the redundant Read Code and moved to the current Read Code.

- **Improper synonyms**. When a synonymous term (e.g. *'Cardiac rupture'*) for a Read Code does not mean the same thing as the preferred term (e.g. *'Myocardial infarct'),* then the synonymous term is moved to a different Read Code.

- **Ambiguity**. If a term is found to be ambiguous (e.g. *'Cord compression'*) and that ambiguity is not resolved by the term already being a synonym to two or more different preferred terms (e.g. *'Umbilical cord compression'* and *'Spinal cord compression'*), then the user is offered two or more Read Codes to choose between. One of these is usually the original concept, unless that has been made extinct because it is the *preferred* term that is ambiguous.

- **Obsolete term.** Terms are occasionally marked as obsolete, for example when they contain a spelling mistake or are a close equivalent to another term (e.g. if both *'Infection in finger'* and *'Infection – finger'* were discovered, then the latter might be made obsolete). In this case, although the relation between the obsolete term and the Read Code is removed from the CTV3 release, no new Read Code is suggested, or needs to be, for analysis.

Classifying changes in these four different ways makes it possible to suggest a strategy for users. In the case of:

- **Obsolete terms**, no change is required.

- **Redundancy**, it is advised that the Read Code for analysis can be updated automatically.

- **Improper synonyms**, it is advised that the user checks that the change is correct in principle, and then it can be applied to update the Read Code used for analysis in all patient cases automatically since there is no ambiguity.

- **Ambiguity**, the user needs to check each patient case to resolve the ambiguity before updating the Read Code used for analysis, unless there are exceptional circumstances. This is potentially time consuming, but experiments with live data have shown that this option occurs in a very small number of suggested changes.

The DCF is now considered an integral part of the CTV3 release product and is of use in the following circumstances:

- As part of the update process following each CTV3 release.

- When migrating from the 4 Byte GP Read Codes, or Read Codes Versions 1 or 2, to CTV3.

- When receiving clinical messages because, when taken together with the Description file, it is a source of all descriptions (relationships between terms and Read Codes) ever released in CTV3. It is therefore a resource to check that incoming Read terms and Read Codes form valid CTV3 descriptions.

Responsibility is shared between three parties. The UK Terminology Centre as the terminology provider needs to supply the DCF as an accurate reflection of changes made in the release file. System developers need to design clinical information systems to be able to handle change. End-users need to respond to the changes, with help from their system suppliers, to ensure accurate results when running queries, reports and for decision support triggers (e.g. drug interactions to be detected). If any party fails to meet its responsibilities, there is a risk that queries for audit, reports and decision support may not produce the desired result. This is because when the query is written, the user sees the current relationship between terms and Read Codes. If these relationships change over time, then the query may fail to retrieve the expected data from the patient database.

The solution is considered **robust** because it provides some protection when parties only partially meet their responsibilities.

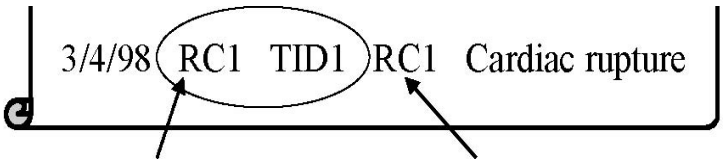
- By maintaining the user's selected Read Code, Term and Term identifier, the clinical record does not change.

- By relating all suggested updates of the analysis Read Code to the end-user's selected Read Code, an anchor is provided which allows changes to be reversed (1) by the UK Terminology Centre in subsequent releases of the DCF if an error in the file is found, and (2) by the end-user if the change subsequently proves undesirable for some local reason.

- By maintaining the set of changes as the best mappings at that point in time, users who have not made all changes after a release can carry on making changes at any point in the future using the most up-to-date DCF without penalty.

Diagram 1: A patient database may contain both the Read Code selected by the user (RC1) and a parallel, sometimes updated (e.g. RC2), Read Code for information retrieval and decision support (the 'Analysis Read Code') as well as the unchanging term identifier (TID1). Note that this diagram depicts a literal interpretation of 'analysis' Read Code and that there are various implementation alternatives.

Clinical record on 31/7/98

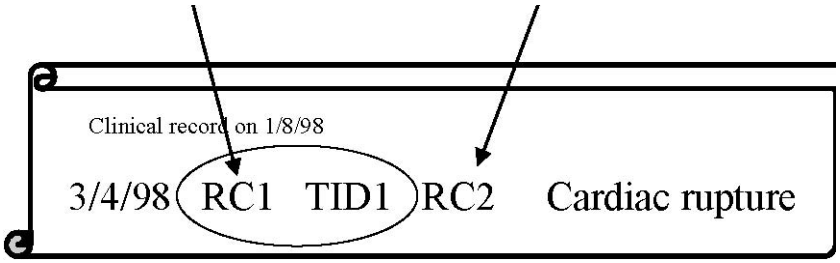


Medicolegally sound. Users selection is not disturbed

Selected description remains unchanged

Analysis Read Code may change over time

Information retrieval improved as changes in CTV3 are reflected in patient data.



Robust. Errors in the description Change File can be corrected

2. Description of this Document

This document describes:

The reasons for change in which terms are associated with which Read Codes in CTV3.

The structure of the DCF which maps deleted descriptions onto new Read Codes to be used for information retrieval.

The principles that are required for handling change safely in patient data based on advice contained in the DCF. The principles reflect good practice and need to be met in any implementation of the file.

Detailed advice, with extensive examples, about how the file will be populated by the UK Terminology Centre.

Implementation advice. These are guidance notes gathered together following discussions with systems developers. They are informative and intended as a source of advice, *not* as a means to control system implementation.

This document does **not** describe how change to qualifiers or the hierarchy should be handled. These issues are less pressing, but are now the subject of new work. Handling change in qualifiers will be covered in a later document, although the design principle of distinguishing ***selected*** and ***analysed*** Read Code will be the same.

3. New Terms in CTV3

Clinical Terms Version 3 contains more terms than previous versions of the Read Codes, covering the needs of many user populations (nursing professions, dieticians, physiotherapists, surgeons, general practitioners, etc.). Accompanying this broad scope is significant change to the content between CTV3 releases. There are two types of change.

- New material is added. Concepts, terms, hierarchical relationships and qualifiers are introduced in response to users' needs.

- Errors and inconsistencies are corrected. Despite having rigorous quality assessment (over 500 validation rules need to be satisfied before each release), a large-scale effort such as CTV3 will inevitably contain some errors. Experience shows that some kinds of error are only likely to be picked up when terms are in active service use.

3.1 Why are new additions required and how do errors occur?

- While some clinical areas have a long tradition in coding and terming schemes, in other areas there has been no previous work. Populating these areas fully and integrating them with other clinical terms takes several cycles of development and use in the real world during which redundant Read Codes (duplicates) may be discovered, along with other types of error.

- As CTV3 is a superset of older versions of the Read Codes (it contains all of the terms and Read Codes from older versions, although 4 Byte Read Codes have an extra leading dot), it must handle a legacy of a decade of coding using these versions. Designed with the constraints of computer hardware of the time, only a limited number of codes and levels in the hierarchy were available to coding authors. Many terms are therefore inappropriately associated with Read Codes whose preferred term means something else. Users migrating to CTV3 will wish to retain the old Read Codes for medico-legal reasons. However, at the same time, they may wish to assign new current Read Codes where appropriate to bring their data more in line with the rest of CTV3.

- During the process of creating CTV3, errors in Read Codes Version 2 were 'corrected' e.g. improper synonyms were re-allocated to new Read Codes in the release files. However, after prolonged discussion, there was general agreement that any CTV3 Read Code that is also found in an earlier version should have its older terms reassigned to it (and marked as an extinct concept). This work is complete, but has meant that users of CTV3 prior to October 1997 have had to (or will need to) reassign some Read Codes in their clinical records now associated with different terms.

- Even in the mature areas of the thesaurus, new developments in medicine and new requirements for clinical information systems prompt a constant stream of new terms for such things as drugs, appliances, operations, and disorders.

3.2 How is change handled in other terminologies?

The traditional classifications, such as ICD, do not allow change except in formal revisions, which come at intervals of greater than a decade. No official mapping provided to support this process and versions are not upwardly compatible.

Older versions of the Read Codes (e.g. 4 Byte set and Versions 1 and 2 of the 5 Byte codes) changed relatively little between releases. While stability is always welcome, the result was that many useful clinical terms did not appear in these versions and in many cases could not be inserted without creating difficulties. Because the file structure only allows five levels of children and around 60 siblings, new concepts were frequently added as improver synonyms. However, doing this may lead to cases being inappropriately retrieved when reports are run. Terms are occasionally marked as deleted and reassigned to other Read Codes, in which case users cannot find them when keying and patients with these Read Codes may subsequently be missed in searches.

More modern terminologies are beginning to look at mechanisms for handling change. In discussions between NHS connecting for Health and other terminology groups, both SNOMED and the National Library of Medicine (Washington, US), who are responsible for the Unified Medical Language System, asked how NHS Connecting for Health handled change and was invited to comment on their suggestions. So far, organisations are concentrating on how to describe how terms can be assigned to new concept identifiers (e.g. Read Codes). The motivation behind describing change is primarily to support integration of terminologies and multi-author creation of terminologies. No other organisation seems to have made a contribution to how descriptions of change can be grouped to support both clinical system designers and also clinicians who must design and use robust information systems.

3.3 What changes occur in the content of CTV3?

In anticipation of the need to handle significant change to its content, the file structure for CTV3 was designed to be flexible. New concepts, terms, descriptions (the way terms are linked to Read Codes), hierarchy relationships, atoms and qualifiers and cross-maps can be (and are) regularly added.

Concepts (Read Codes) are never deleted. They may be marked as ‘Extinct’ if there is a significant problem with the concept, or ‘Optional’ if there is a problem, but some community of users still needs to use the concept in some circumstances, or ‘Redundant’ if a concept is found to be a duplicate of another.

Terms containing spelling mistakes are being corrected, but all other term corrections are severely restricted. Every correction is checked by two authors to ensure there is no change in meaning (see Appendix 1 for a list of allowed changes to terms). The UK Terminology Centre plans to stop authoring any corrections to terms completely in the future. At that point, terms containing errors will simply be marked as obsolete and detached from their Read Codes in the Description file. New, correct terms will be introduced at the same time.

Descriptions (the link between terms and Read codes) may be deleted. The rule is that terms may only be synonyms to the preferred term for a Read code if the one of the meanings of the synonymous terms is the same as that of the preferred term. Coding authors delete descriptions that so not fulfil this criterion.

Hierarchy relationships can be deleted. If a child is found not to be a 'kind of' its parent, then the link is removed. For example, in earlier versions of Read Codes, '*Constipation*' has not '*Not constipated*' as a child. If this parent-child relationship had been erroneously introduced into CTV3, then it would be deleted. This ensures that searches for all patients with constipation will act as a user will intuitively expect.

Atoms and qualifiers may be deleted, or one of their flags amended. If an atom (e.g. *Site: Hip joint*), which defines some aspect of the meaning of a concept (e.g. *Osteoarthritis of hip*) is found to be false, or if a qualifier (e.g. *Priority: Urgent*) is not applicable, then these are deleted from the template file. Extensive quality assurance reduces the chances of this happening. For example, the procedure and disorder parts of the hierarchy are checked automatically against the anatomy hierarchy to make sure the two are similarly structured (if *hip joint* is a type of *joint*, then *osteoarthritis of the hip* should be a type of *osteoarthritis*).
Cross-mappings that are found to be incorrect are deleted. These have been extensively validated and further tested in real use, but there is still a trickle of errors which are being corrected when they are detected.

3.4 Change Files versus Incremental Files

Two kinds of change file are required. These are called **Incremental files** and **Change files**. **Incremental files** list the additions and deletions to each release file since the previous release. They are designed:

- To support upgrading of the terminology resource files (CTV3 look-up files) for clinical information systems.

Implementing incremental files has the same affect of substituting one set of release files for a previous outdated set of files. Therefore, it affects picking lists that the user sees when entering or selecting data to go into a patient record, or when devising a new query. It also affects the way that previously written queries will run (see later for analysis of this). ***The UK Terminology Centre is looking separately at the issue of publishing incremental release files. This issue is not examined further in this document***. Each release file would have a shadow incremental release file. Incremental changes will need to be applied in a strict order.

Change files list corrections that might be made to the Read Codes used for analysis in patient records, and, potentially, in queries stored in clinical information systems. Their function is therefore:

- to support search and reporting of the Read Codes in the patient records

As we will see later, a change file is an up-to-date listing of the best Read Codes to use for analysis for any description that has been changed. Therefore, changes may be applied in any order. Users will only need access to the latest change file. If there has been insufficient time to make all of the recommended changes prior to a new DCF being released, users may immediately discard the old DCF and continue with the new version.
Four areas of CTV3 need to be covered by change handling mechanisms. These are:

- changes to descriptions

- changes to qualifiers and atoms (semantic definitions)
- changes to the hierarchy
- changes to mapping files.

Change in the Description file, the first of these areas, is considered to be a priority area and the rest of this document is limited to a discussion of the management of this process. The DCF has one subsidiary function, in addition to the general function of change files listed above. This is:

- To act as a repository for deleted descriptions, which users may wish to check when receiving data from electronic sources.

3.5 Managing Change in CTV3

To manage change, it is important to reduce the need for change in the first place, and to restrict the effects of change. Each of the following plays an important part.

- A formal mechanism for change control in CTV3. For example, Read Codes are never deleted and their preferred terms for a Read Code are only changed according to strict rules. This is allowed in situations such as when the preferred term is found to be ambiguous (and then the concept is marked as extinct), or in the case of items such as drugs, when standard naming conventions are changed.
- A formal mechanism of quality control. The UK Terminology Centre now has more than 500 quality assessment rules in place. Quality assessment routines are run every night so that errors can be corrected immediately. Two authors inspect changes in critical areas when the changes cannot be automatically verified.
- A formal mechanism for testing new updates. The UK Terminology Centre has concentrated recent effort on reacting to pilot sites which discover important and relevant omissions, terms attached to a Read Code which are not true synonyms, and hierarchy misplacements. Future change should be far more limited as a result.
- A standard mechanism for alerting system developers to changes that need to be made. The DCF described in this document is one vehicle for doing this
- Safe handling of any changes that need to be made in the clinical record to ensure that these do not affect its ability to support clinical care or to act as a medico-legal record.

4. Types of Anomaly leading to changes in the Description File

Terms are re-allocated to new Read Codes when any of the following are found:

1. **Improper Synonyms**. A term is found not to share the same meaning as the preferred term of a concept.

    - For example, when *Cardiac rupture* was found to be a synonym of *Myocardial infarction*. The preferred term (*Myocardial infarction*) remains attached to the Read Code, but the improper synonym (*Cardiac rupture*) is either moved to a new Read Code (as described below) or to a pre-existing Read Code if this has the same meaning.

    Former release

Read Code	Term	Status of term
RC1	Myocardial infarction	Preferred term
	Cardiac rupture	Synonymous term

    Current release

Read Code	Term	Status of term
RC1	Myocardial infarction	Preferred term
RC2	Cardiac rupture	Preferred term

The new Read Code will be of the same Subject type[1] as the old Read code and will usually be of the same medical type. Occasionally disorders may be moved to the 'History and observations' chapter (same subject type, but different medical type), but the intention is to limit these to a minimum.

2. **Ambiguous term.** A term is found to be ambiguous.

    - For example, if the term *Fit* was the preferred term of a concept, then it may mean '*Healthy and well'* or a '*Convulsion'*. As the meaning of a concept is determined by the meaning of its preferred term, this is unacceptable. The concept (Read Code) is made extinct and given a new preferred term which captures both of the old meanings (e.g. '*Convulsion or healthy and well'*). The two or more meanings are made clear by assigning the ambiguous term as a synonym to two other concepts (Read Codes), each with an unambiguous preferred term. These other Read Codes may already exist in the terminology or may be newly introduced.

    Former release

Read Code	Term	Type of term
RC1	Fit	Preferred term

[1] Medical types are largely equivalent to the chapter headings in CTV3, i.e. the concepts sitting just below Read thesaurus, the top node in the Read hierarchy. Subject types are generally broader than medical types (all low level 'History and observations', 'Test results' and 'Disorders' share the same subject type), but distinguish a change in type. They only play a role in the 'History and observations' chapter where findings categories (sodium level) are distinguished from potential findings (low sodium level). The latter sit below the former in the hierarchy.

Current release

Read Code	Term	Type of term
RC1	Convulsion or healthy and well	Preferred term
	Fit	Synonym
RC2	Healthy and well	Preferred term
	Fit	Synonym
RC3	Convulsion	Preferred term
	Fit	Synonym

RC1 is marked as extinct. RC2 and RC3 are introduced as current concepts. Usually the new Read Codes will both be of the same subject type and medical type as the original concept. Occasionally, however, the reason for a concept being ambiguous is that it can be interpreted either as a procedure or a finding. In this case, one of the alternatives will be of a different subject (and therefore medical) type to the Read Code of the concept with the ambiguous preferred term.

In other cases, a synonym is found to be ambiguous. This is allowed in the case when the term (e.g. '*Cord compression'*) appears against different Read Codes whose preferred terms (e.g. *'Umbilical cord compression'* and *'Spinal cord compression'*) make the distinction clear. However, if one of the meanings is not associated with a Read Code, then a new Read Code is introduced, and the synonym is mapped to it in the Description file. Now users are offered the old and new Read Codes in the DCF.

3. **Redundant concept**. Two terms, attached to different Read Codes, are found to refer to the same concept.

   - For example, if '*Myocardial infarct'* and '*Myocardial infarction'* had different Read Codes, then the Read Code for '*Myocardial infarction'* may be chosen as the current code, and the terms associated with '*Myocardial infarct'* in the Description file moved to be associated with the current Read Code. Finally the Read Code previously linked to the term '*Myocardial infarct*' would be marked as redundant.

4. **Obsolete term**[2]**.** A term e.g. '*Bronicle asthma*' is made obsolete when it is considered not to have any use in recording clinical notes. The following decision algorithm is used to amend the CTV3 release files:

   - If the meaning of the concept behind the obsolete term is clear and the term to be made obsolete is a synonym, then the term is flagged as obsolete and removed from being a description of the concept.

   - If the meaning of the concept behind the obsolete term is clear, but the term to be made obsolete was the preferred term, then this old preferred term is flagged as obsolete,

[2] Note that obsolete terms are different from what used to be called obsolete concepts (now optional concepts). These are concepts which are deprecated for some reason, therefore best avoided, but are still required for some purposes.

removed from the Description file and a new preferred term linked to the concept (e.g. '*Bronchial asthma*').

- If the meaning of a preferred term is unclear and the meaning of the concept is unclear, then the concept is made extinct and the term left as current (and therefore not marked as obsolete).

5. **Concept out of step with previous version of Read Codes** (i.e. to ensure that CTV3 is a true superset of previous Read Code versions). As part of the work of creating the CTV3 superset, all of terms that are associated with Version 2 Read Codes[3] now have the same codes in CTV3. Originally these codes had been 'cleaned up' to make them compatible with CTV3 principles and improper synonyms removed (for example). The 'cleaned up' concepts have been given new CTV3-only Read Codes. Note that these changes only affect people who used CTV3 prior to October 1997, in keeping with the principles of CTV3. Further details and examples are given in the document "The Read Codes: Update for Read Codes Version 3 users of the March 1998 release (Superset)" which is available from the UK Terminology Centre. These are not discussed in further detail in this document.

5. Implications for Users of Changing Descriptions

Clinical information systems should store the Read Code, term identifier and term that a user has chosen. Following term re-allocation, the term and term identifier remain unchanged and so the clinical record will look the same. However, the Read Code stored in the clinical record may not match the Read Code against the term in the Description file in the latest release after a term reallocation. This can affect searches across patient records and decision support. There are implications both in re-allocating terms to new Read Codes and in not doing this. In the examples below, it is assumed that searches are based on the Read Code (and not the term identifier or Read Code plus term identifier)

If the UK Terminology Centre fails to correct an error, or system developers fail to implement

some mechanism for change control, or users do not invest time to make changes in those cases where their input is needed, then the following errors may occur:

Improper synonym flag	Failing to respond to these will mean that a search (e.g. for all patients with cardiac rupture) will bring back patients who do not match the search criterion (e.g. all patients with a myocardial infarct – fortunately not everyone with a heart attack ends up with a cardiac rupture).
Ambiguous term flag	Failing to respond to these will mean that a search (e.g. for all patients with a 'fit') will bring back patients who do not match the intended search criterion (e.g. patients who are fit and well).
Redundancy flag	Cases not matching the user's query code (if this contains just one of the codes) will be missed.

[3] 4-byte Read Codes were added later to CTV3, simply by adding a preceding dot. None of the 4-Byte codes is marked as current in CTV3 and so none have had to be re-allocated.

Obsolete term flag	Failing to respond to these will not affect searching and reporting. In fact, unless making the term obsolete is combined with some other error, the mapping in the DCF is to the same Read Code as the user selected.

If the UK Terminology Centre makes changes to CTV3 files and users make changes to the **Read Codes in patient records** that are searched in response, but fail to amend **queries** which were written prior to the change, then the following errors may occur:

Improper synonym flag	The query may expect to find a term against an inappropriate Read Code, but the Read Code has been changed in the patient data, so the patient is missed.
Ambiguous term flag	If a preferred term is found to be ambiguous, then the concept will be made extinct. A search based on this concept will not pick up any of the children previously found underneath the code in the hierarchy. Cases will be missed. A query from some point higher in the hierarchy will also fail to retrieve the patient, as the extinct concept may have been moved out of that part of the hierarchy to a more suitable position (better reflecting its combined meaning).
Redundancy flag	A query based on a code that has become redundant may miss all of its old children (these are now children of the current concept).
Obsolete term flag	No effect on search.

From a practical standpoint, new searches and reports are improved by term reassignments in CTV3 and the clinical record. But any one old search may be improved or disadvantaged by a term reassignment.

- For example, when a search is run to detect all patients with chronic respiratory disease (in order, for example, to reach a priority group of patients needing influenza immunisation), it is an advantage to have each of the diseases in the hierarchy under chronic respiratory disease defined as correctly as possible.

- However when looking at a management trend, for example when comparing service activity between quarters, change in figures without an underlying real change is a disadvantage.

There are ways of circumventing the disadvantages while retaining the advantages. One is to re-run the report on the old data using the new Read Code release and reassigned terms in clinical records. Then the comparison is based on common ground.

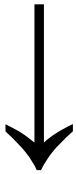
6. Principles of Changing Read Codes in Patient Records

To safely correct improper synonyms, ambiguous terms and newly discovered redundant codes, the following principles are proposed. They are based on the idea that it is possible to

distinguish between the Read Code selected by the user and the Read Code used for searching and reporting. When first selected, these two Read Codes are the same. However, when errors are found in the Read terminology or new clinical knowledge forces a change in perspective, a new Read Code may be substituted for searching and reporting although the selected Read Code, the term, and term identifier remain unchanged.

Original clinical record file

Selected RC	Analysis RC	TermID	Term
RC100	RC100	Term1	Cardiac rupture
…	…	...	...



Description change file

Changed clinical record file

Selected RC	Analysis RC	TermID	Term
RC100	RC978	Term1	Cardiac rupture
…	…	...	...

A. Any update to the clinical record should neither adversely affect the care of a patient, nor the medico-legal status of the clinical information system.

The Read Codes, term identifiers and terms selected by a user to record the process of clinical care should be stored in a clinical information system and should not be amended. Change to descriptions can be seen as one of three types of changed record entry, all of which should ensure that the original details are not deleted and, at minimum, are stored in an accessible and verifiable audit trail.

1. The original user made a true error. For example, '***Localisation of placenta'*** was chosen in the heat of a consultation by one user who intended to record that an ear swab had been taken. This was a clear error, which the physician was quite unable to explain, and underlined by the fact that the patient was male. When this error was pointed out to the clinician, he immediately wished to change the entry.
2. A clinical detail is superseded. A patient is finally shown not to have '***Asthma'***, which had been the previous diagnosis.
3. A clinical term is allocated to a new Read Code for information retrieval. It is essential that the original selected Read Code can be seen by the user, and can be used by the system for mapping in the future. It is only this type of change that is the subject of the paper.

B. Information retrieval for decision support, audit and reporting needs to take account of terms being moved to different Read Codes as new knowledge arises or when correcting errors.

To do this safely, system developers need to distinguish between the user's originally selected Read Code and the Read Code used for searching and reporting (the analysis Read Code). Mechanisms[4] that may be used for this purpose include (but are not limited to) the following:

- Creating and maintaining mapped versions of the recorded code within, or linked to, the record (e.g. introducing an 'analysis Read Code' field in an event record). This literal rendition of the principle is used in examples throughout the document.

- Use of mapping tables to include patient records (containing only selected Read Codes) in an appropriate way when searching for, or retrieving, information. The query is updated in the light of information about changes to descriptions.

- Applying the mappings to modify records in a manner that is subject to a versioning process, sufficient for medico-legal audit, that indicates the source of the change and facilitates roll-back to the pre-mapped version. In this case the selected code, now part of the version history, must always be available for the user to inspect and for the system to access if further changes are suggested at some time in the future.

C. Users of CTV3 should be able to automatically update their patient records (or queries) when receiving each new release whenever this can be done safely.

Automatic update is important because user intervention is time-consuming.

Read Codes found to be **redundant** (e.g. duplicates) are a case where there is no advantage to manual intervention. Automatic update should be seriously considered.

At the other end of the spectrum, where terms for Read Codes are found to be **ambiguous**, more than one Read Code will be offered to the user. Changing the Read Code used for searching and reporting requires user intervention to confirm which of the potential meanings was intended. In the case of a term like '*Cord compression*,' orthopaedics surgeons, using a system designed for them, may know that all patients with this term identifier-Read Code combination had '*Spinal cord compression'* (rather than '*Umbilical cord compression'*, which happens to new-born babies). However, a general practitioner could not make this assumption and individual patient records would need to be checked to ensure the correct alternative is selected. However, a safe automatic re-allocation will also be provided, though the only safe alternative may be the original ambiguous Read Code. While resolving ambiguity typically requires the user to check each patient record, implementing the DCF on a real patient database has shown that these form a small minority of changes in real databases.

In the case of **improper synonyms**, users need to check the list of proposed reassignments but, as the terms are not ambiguous, it is not clear that individual patient records need to be checked. Therefore, if the suggested re-allocation is acceptable, the Read Code used for analysis in all patient records containing these descriptions may be

[4] These options are discussed in more detail in the Appendix on implementation guidance.

updated automatically.

When terms are made **obsolete,** the Read Code in patient records does not change. It is not necessary for users to change the term as searching and reporting will be unaffected, and it is felt unlikely that users will wish to spend effort doing this. If they wish to do this, e.g. to correct spelling mistakes, then this should be done on a case-by-case basis. Changing terms is discouraged unless there is a system function for seeing the term that has been superseded. Systems must keep an audit trail as a minimum precaution if term changes are to be allowed. Although the UK Terminology Centre plans to build a mapping file from obsolete terms to new terms where there is an obvious match, the current plans only include using this as a resource for maintaining the DCF.

The case of concept re-allocation is special. Concept re-allocation is a one-off process. The list of these was provided in October 1997 and will not be expanded. Only users of CTV3 prior to this release need to implement concept re-allocation. However, as there is potential for confusion between the meaning of Read Codes before and after re-allocation, then in this case only users are advised to change the selected code as well as the analysis code. A record of the patient record prior to concept re-allocation must be archived. The UK Terminology Centre has contacted all known suppliers in this category. Those users who need to do concept re-allocation are strongly advised to discuss their proposals first with the UK Terminology Centre.

Summary of recommended user responses

Improper synonym flag	Users need to check they are happy with the proposed list of changes to the Read Codes used for analysis. For all those users that are happy, automatically substitute the newly suggested Read Code in the analysis field.
Ambiguous term flag	Flag all patient record entries containing ambiguous terms, so that the user can resolve these on a case-by-case basis at a suitable time.
Redundant concept flag	Automatically change the Read Code in the analysis field to the new Read Code, which will be the current Read Code for that concept.
Obsolete term flag	The Read Code will remain unchanged. It is not therefore necessary for users to change the term when searching and reporting (it is felt unlikely that users will wish to spend effort doing this, but if they wish to do so, e.g. to correct spelling mistakes, then this should be done on a case-by-case basis).
Concept reallocation flag	Automatically change the selected Read Code. The old selected code should be archived first. This is a one-off change for CTV3 users prior to October 1997 only. Users are strongly advised to discuss with the UK Terminology Centre prior to making the change.

D. Users should be able to recover from wrong changes.

The UK Terminology Centre should work with developers and users to protect users from any errors made when advising changes that are subsequently withdrawn.

Experience suggests that better re-allocations may only be found after extensive testing in use. Therefore, another reason for systems retaining the original Read Code and term identifier selected by the clinician is that these can be used in subsequent re-allocations of the Read Code used for searching and reporting.

E. Sequences of changes should be handled robustly.

Sometimes more than one change may affect a description. An improper synonym may be moved to another Read Code which is subsequently discovered to be a duplicate (redundant). Alternatively, a previously suggested re-allocation may have to be corrected or withdrawn. A system which depends upon the user applying a sequence of changes in the correct order is likely to be more error-prone. Therefore, rather than being a sequence of changes, the DCF contains the current best mapping(s) from each term identifier and Read Code that a clinician may have used. The newly-listed best mapping will be valid whether or not the initial change was made.

F. When communicating data coded using CTV3, sufficient information should be transferred to allow receiving users to retrieve that information usefully once that information is stored in their patient record system.

Current messages (e.g. pathology links) use defined populations of Read Codes, when the rules for transmitting and receiving Read Codes are simpler. Future clinical messages, such as hospital discharge summaries or exchange of records between GPs, may include more general coded clinical information. In this case, the UK Terminology Centre would strongly suggest that the **selected** Read Code, term identifier and Term rubric are included in the message. Including the **analysis** code too may disambiguate ambiguous descriptions, which only the sender may be able to do.

Electronic messages may come from systems that have a long legacy of usage and include many descriptions which are no longer in CTV3. These may be mapped on to new analysis codes by the receiver to better integrate the incoming data with other system's information for retrieval purposes.

Electronic messages might contain local codes and terms (though this should not happen). As the DCF, together with the Description file, contain a complete list of descriptions ever used in CTV3, receiving systems may check to see if incoming descriptions are nationally valid. Descriptions involving local codes and local terms should not be stored in the receiving system's Read coded patient database unless recorded.

7. Structure of the DCF

An entry is made in the DCF when a Read term (V3_TERM_ID) attached to a Read Code (READ_CODE_PREV) is either allocated to a new Read Code (READ_CODE_NOW) or made obsolete, in which case READ_CODE_PREV = READ_CODE_NOW. The date of the release (RELEASE) in which the change was made is recorded using ISO 861 format YYYY-MM-DD. For example, the first day of October 1997 is recorded as 1997-10-01.

7.1 DCF File Structure (dcf.v3)

FIELD NUMBER	TITLE	SIZE
1	V3_TERM_ID	Char(5)
2	READ_CODE_PREV	Char(5)
3	READ_CODE NOW	Char(5)
4	MAP_STATUS	Char(1)
5	RELEASE	Char(10)

7.2 Map Status Flags

The mapping is also marked with a MAP_STATUS which can be 'C' (Concept re-allocation), 'R' (Redundant change), 'S' (Improper synonym), 'A' (Ambiguous term), or 'O' (Obsolete term).

C	**Concept re-allocation**. ***This flag will not be found in standard releases of the DCF***. Concept re-allocations are only required by users of CTV3 prior to October 1997 and these are released in a one-off file, containing just C flagged rows, available on request from the UK Terminology Centre. The contents of this release of the DCF will exactly mirror those of the Concept Re-allocation File, but in DCF file format for those who only wish to write a load programme for a single 'changes' file structure. Users of CTV3 systems collecting data with CTV3 prior to October 1997 must re-allocate concepts using the DCF or CRF before implementing the rest of the DCF functionality. The C flag is therefore disregarded in the analyses presented below.
R	**Redundant change**. Only term re-allocations in which the Read Code is made redundant are marked as 'R'. Users are advised to change the Read Code used for analysis.
S	**Improper Synonym**. The user is advised to change the Read Code used for analysis to the new code. No alternatives are offered in this situation (though the user may choose to stay with the selected code).
A	**Ambiguous term**. As the meaning of the term is ambiguous, alternative Read Codes are provided to the user. The alternatives disambiguate the meaning.
O	**Obsolete term**. In these cases *READ_CODE_PREV* = *READ_CODE_NOW*. They are included to fulfil the principle that all descriptions that have been valid at some previous point in time are included in the DCF. Note that the change file does not provide a mapping to a new TERM_ID.

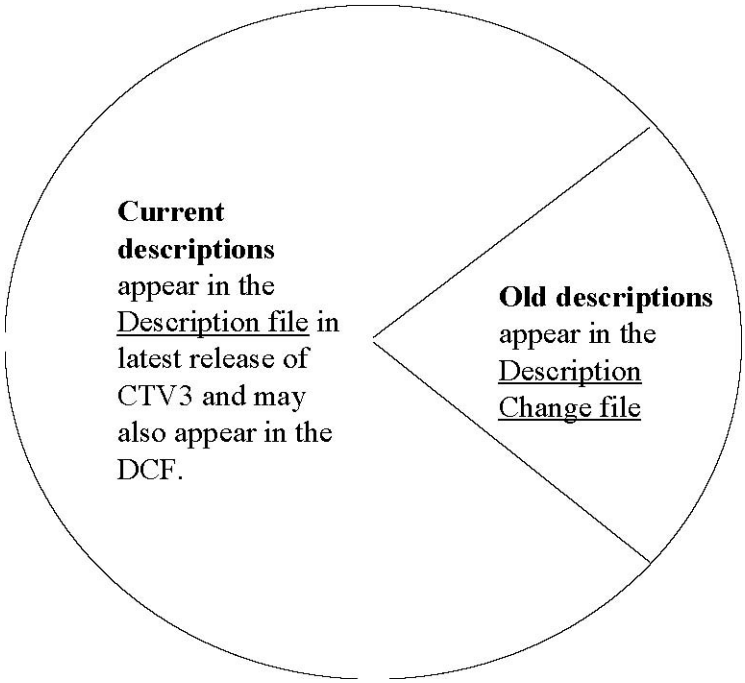
The DCF does not supersede either the 4 Byte Read Code and Term Mapping File or the Version 2 Read Code and Term Code Mapping File. These are still required for migration, as they provide the term identifier required for CTV3. The 4 Byte Read Code and Term Mapping File maps all terms and Read Codes in the 4 Byte set onto valid CTV3 codes and term identifiers. Likewise, the Version 2 Read Code and Term Code Mapping File maps Version 2 Read Codes and term codes onto CTV3 Read Codes and term identifiers.

If the DCF is used at each release, then the Redundant Codes Mapping File can be ignored. This simplifies the procedure for searching and reporting on Read Codes.

8. Contents of the DCF

The purpose of the DCF (filename dcf.v3) is to provide the most up-to-date assignment of a Read Code to each term. The file contains every description (as READ_CODE_PREV and V3_TERM_ID combination) that has been released in CTV3 after October 1997 and that no longer appears in the Description file (see Diagram 2). For the period before October 1997, only old descriptions known to have been used in live CTV3 sites will be included. NHS Connecting for Health has had discussions with all known CTV3 user sites about this during 1997/1998.

The file will be updated at each release. As CTV3 is a superset of older version of the Read Codes, the DCF also contains descriptions that have been removed from the 4-Byte GP set and from Versions 1 and 2, with mappings to the best available CTV3 concept.
Diagram 2: all CTV3 descriptions that have been released after October 1997 (or which are known to have been used prior to this date) appear in either the current Description file or the Description Change File, or both.



8.1 The DCF contains descriptions removed from CTV3

The following kinds of description have been removed from CTV3, and are therefore deleted from the Description file and placed in dcf.v3.

- Descriptions involving Read Codes made **redundant**.
- Descriptions involving terms which are found to be **ambiguous**.
- Descriptions involving terms found to be **improper synonyms** of their preferred term.
- Descriptions involving terms that have been made **obsolete**. In these cases, *READ_CODE_NOW* will be the same as *READ_CODE_PREV*, unless the term is

subsequently re-allocated to a new concept. An example of this is given later.

- In addition, a separate release of the file is available containing only descriptions that have been reassigned when building the Superset (**concept re-allocation**). This release is only required by users of CTV3 prior to October 1997.

8.2 The DCF is about core concepts.

Core concepts are one of two kinds of concept that exist in CTV3. Core concepts include disorders; procedures; signs and symptoms; personal, social and family history; tests; results, drugs and appliances. Core concepts may be directly selected by users (though note that drugs need to be selected in an action context, such as prescribing).

Other concepts are intended for use as attributes (e.g. '*Laterality'*) and values (e.g. '*Left'*), qualifying the core concepts (e.g. '*Osteoarthritis'*). This second group contains many simple descriptive phrases that may be ambiguous (e.g. with preferred terms such as '*Right'*, '*Left'* and *'Flat'*). However these are intended for use with attributes that clearly describes their relationship with core concepts e.g. '*Fracture of hip - Laterality: Left'*. In this context, 'Left' has no political connotations. To correct such ambiguity would force users to disambiguate contexts which are clear from the expected use of the terms.

Handling changes to attributes and values needs to be considered within the context of changes to templates, which are beyond the scope of this document

8.3 The DCF may contain alternative mappings

For any one combination of previous Read Code and term identifier, there may be several alternative mappings. The following are possible:

- One record with an R flag
- One record with an O flag
- One record with an S flag
- Two or more records with an A Flag
- One record with an R flag and one record with an S flag
- One record with an R flag and two or more records with an A flag

Alternatives occur when:

- An ambiguous term is found and can only be resolved by choosing between two or more concepts, one of which may be the original concept.
- One change (a Read Code is found to be redundant) is followed by another (e.g. a term attached to the redundant Read Code is found to be an improper synonym, or a term is found to be ambiguous). Leaving in the redundant mapping always allows users to implement minimum change (correcting redundancy), or to alter an unsatisfactory change back to the current Read Code.

The range of possible scenarios in which re-allocations are combined is provided in a later section.

8.4 The DCF is a snapshot containing the current best mappings of old descriptions

With each release of CTV3, the DCF is amended to include only the current best set of mappings from old descriptions to new ones. It is not a cumulative file nor an incremental log, and therefore if users do not apply one version of the DCF before another comes out, they may immediately move on to using the new file. Similarly, if users have made some changes by the time a new DCF appears, but not, for example, corrected all ambiguity in their patient databases, they may immediately move onto using the next version of the file without any penalty.

9. Using the DCF

The DCF will be useful in several circumstances. These include:

- Upgrading clinical records with each new CTV3 release
- Migrating from the 4 byte GP set
- Migrating from Clinical terms Version 1 or 2
- Accepting Read Codes from a clinical message

The requirements of each of these circumstances are discussed below. Common to all of these is the following suggested response. It assumes that the user has looked at a list of improper synonyms with their old Read Code plus preferred terms and newly proposed Read Code plus preferred terms, and indicated which of these can be automatically changed in patient records.

For clarity, the advice is given as though the patient records actually contain an analysis field. Some system developers may prefer a 'virtual' analysis field[5], constructed when querying the database. Although the principles in the algorithm and the rest of the advice in this section remain true, the implementation will be different.

Status	Action
If there is a single entry (record) in the and READ_CODE_PREV, with a rele DCF:	DCF for a combination of V3_TERM_ID ase date after last application of the
and map status = R	For each patient records: Automatically assign new analysis Read Code in patient records.
and map status = O	Can ignore as READ_CODE_PREV = READ_CODE_NOW.
and map status = S	Check clinician is happy in principle with re-allocation of this Read Code. Check for any dependent reports. If neither is a problem then: For each patient record: Automatically assign new analysis code in patient records.

Status	Action
If there are multiple entries (records) in the DCF for a combination of V3_TERM_ID and READ_CODE_PREV, at least one of which has a release date after the last application of the DCF:	

[5] See Appendix B for implementation guidance.

and all have map status = A	User must decide on a case-by-case basis which of the new Read Codes should be substituted for analysis, unless there is a convincing argument to the contrary. Only in the absence of any doubt can a mapping of 'A' be treated like a mapping of 'S'. Even then clinicians should ideally check a report of records changed.
and one record = R, one record = S	For each patient record: If analysis field already contains the Read Code flagged S, no change. Else, if clinicians are happy in principle with the suggested Read Code, assign new analysis code. Else (if clinicians are not happy in principle with the suggested Read Code), and record is already assigned to 'R', no change. Else (if clinicians are not happy in principle with the suggested Read Code, and record is not assigned to 'R') assign to 'R'.
and one record = R, more than one = A	For each patient record: If analysis Read Code is already assigned to one of the codes flagged 'A', and all of the ambiguous alternatives were considered previously (see below), then no change. Else, if already assigned to one of the 'A' group, but not all of the ambiguous alternatives were considered, then flag record to show that user needs to reconsider the alternatives at a suitable time. Else if none of the 'A' group has been assigned, assign to 'R' and flag record to show that user needs to consider the alternatives at a suitable time.

Why does flag ambiguity have to be resolved? Ambiguity needs to be resolved on a case-by-case basis. Therefore in many cases, not all ambiguity will have been resolved by the next automatic application of the DCF. One way of handling this asynchronicity is to flag records as having ambiguity that needs to be resolved at an appropriate time. Then, when a user makes a decision either to resolve the ambiguity by choosing a new analysis code or by actively deciding to stay with the old code (despite it being an out-of-date description). The flag can be reset to 'Decided' status. This way the user is not asked to resolve ambiguity unnecessarily.

When should ambiguity be resolved? This may depend on user preferences. In some cases, many patient records will contain the ambiguous term. Users may wish to resolve ambiguity in a batch session. Others may prefer to make changes next time the patient's record is consulted.

Some may wish to make changes only when reports or decision support are affected.

What information needs to be included in change records? Apart from a field for the new analysis Read Code, system developers will need to consider adding the following fields to each change record.

- The date when the change was made and the name or identifier of the user making the change (for audit trail purposes).

- The CTV3 release date whose DCF was consulted.

- A flag to say whether the change was made:

  - manually (e.g. ambiguity, when the individual case details have been checked),
  - semi-automatically (e.g. improper synonyms where the user has checked the mapping, but not looked at this patient case), or
  - automatically (e.g. redundancy).

- A flag to say whether there is new change to be considered (ambiguity) or not. This may be the first element inserted into a new change record, before any change has been made.

The next time the DCF is applied automatically, it will only be necessary to apply those entries with a release date following the release date of the last application.

9.1 Upgrading CTV3 Resource with each Release

Each Version release will contain the standard release files and the change file(s). Users should first substitute the updated standard CTV3 release files (i.e. the Concept file, Terms file, Descriptions file, Hierarchy file, Template file and Mapping files) for the old versions. This may require a transformation of the contents into the user's format.

Next, Read Codes used for analysis in the clinical record should be substituted by new Read Codes where the DCF advises of changes. The mapping should ignore the old analysis code, and work from the original Read Code and term identifier selected by the user, which must be left unchanged by the process. Mappings should be applied according to the general advice given above. After the process is completed:

- Terms and term identifiers are unchanged.
- The user's original Read Code is unchanged.
- The Read Code used for searching and reporting is updated.

If the DCF has been previously applied, then only those rows with a date following the last release applied need to be processed. However, there may often be case-specific ambiguous alternatives to resolve.
System developers and users need to consider whether queries need to be altered. These include trigger codes for decision support, as well as reports. Developers of Read Code software may consider advising the user of any changes to the synonyms or children of concepts that form

part of the user's stored query. The DCF can be used as a resource to do this.

Finally system developers need to be aware of the need to alter any other part of an implementation that may use Read Codes, such as the specific design of user interfaces. This may be most easily achieved by making all Read coded elements of a system design 'soft' rather than 'hard' coded.

If users regularly update analysis Read Codes in the clinical record using the DCF, there is no need to use the Redundant Codes Mapping file.

9.2 Migrating from the GP 4 byte Read Code set

First, a term identifier should be added to each clinical entry. The 4-Byte Read Code and Term Mapping file (fbtermv3.v3 on the release CD) maps 4-Byte Read Codes and term rubrics onto a valid V3 term identifier.

4-Byte Read Code	Term rubric	CTV3 Term identifier	Date
H14.	Acute tonsillitis	Y02WG	1998-03-01

In some cases, adding a term identifier will be difficult, perhaps because the system supplier has allowed terms to be edited or overwritten by clinical users. However, if term identifiers can be added, users will then be able to use the DCF.

The system application should then add a leading 'dot' to all 4 byte Read Codes. They are then valid CTV3 Read Codes.

Finally the DCF should be applied. Read Codes used for searching and reporting can be changed to the most up-to-date mappings. While it is not necessary to apply the DCF, users doing this will gain significantly in accuracy of searching and reporting because 4 byte synonyms often mean something different from their preferred terms. The problem is that a significant amount of user intervention is required, according to the rules previously described.

9.3 Migrating from Unified Version 2 of the Read Codes

Version 2 Read Codes remain valid in CTV3. Therefore, the Read Code in each clinical entry remains the same when migrating from Read Codes Version 2 to CTV3.

The term stored by the user should not be changed.

However, a term identifier should be added to each clinical entry. The Version 2 Read Code to CTV3 term identifier table (v2termv3.v3 on the release CD) is a resource provided by the NHS Connecting for Health that maps (Version 2) Read Code and Version 2 term codes on to CTV3 term identifiers.

Finally, the DCF should be applied. Read Codes in the clinical record used for analysis can then be changed to the most up-to-date mappings according to the general rules described in this document. While it is not necessary to apply the DCF, users doing this will gain significantly in accuracy of searching and reporting because Version 2 synonyms often mean something different from their preferred terms. The problem is that a significant amount of user intervention is required, according to the rules previously described.

9.4 Migrating from Unified Version 1 of the Read Codes

The unified Version 1 Read coding system contains the same 5-byte Read Codes and rubrics as unified Version 2 but does not contain any term codes. The procedure for migrating from Version 1 to CTV3 is the same as migrating from Version 2 except that when using v2termv3.v3, users must look up on Read Code and term rubric rather than Read Code and term code.

9.5 Accepting Read Codes from a Clinical Message

All Read Codes used in clinical messages are CTV3 compatible. Current messages use small, well-defined populations of Read Codes. However future clinical messages can be expected to use the broad range of Read Codes. The user supplying the data in the message may not be using an up-to date release of Read Codes (whether CTV3 or earlier 4-Byte or Version 2) or may be a CTV3 user who has not applied the DCF. Therefore the system receiving the message should apply the DCF to ensure the most up-to-date Read Code and term identifier combination is used for searching and reporting.
As all Read Code descriptions can either be found in the Description file (if ‘current’) or in the DCF (if ‘obsolete’), the file may be used to check that incoming data is ‘legal’.

10. Detailed Description of how the DCF will be populated

This section describes how the various changes are represented in the DCF. Combinations of changes are also described in detail in this section, so that clinical users and system designers can see how the UK Terminology Centre will react when these occur.

10.1 Improper Synonym followed by further Improper Synonym

If an improper synonym is resolved by re-allocating Term1 from RC1 to RC2 at Date1, the DCF will contain the following:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	S	Date1

At Date2, it is realised that the new combination of RC2 and Term1 is also an impure synonym, so Term1 is re-allocated to RC3. The DCF will now appear as shown below. The original line has been deleted from the DCF.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC3	S	Date2
Term1	RC2	RC3	S	Date2

Note how up-to-date analysis Read Codes are provided for all previous descriptions (both Term1/RC1 and Term1/RC2) in the DCF.

10.2 Withdrawing a Previous Term Re-allocation

If an improper synonym is resolved by re-allocating Term1 from RC1 to RC2 at Date1, then the DCF would be as follows at Date1:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	S	Date1

If *subsequently*, it is realised that an error was made, and that the synonym should have been reallocated to RC3 instead, the DCF would be as follows as Date2:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC3	S	Date2
Term1	RC2	RC3	S	Date2

If instead it was realised that the original re-allocation at Date1 was wrong in the sense that it was unnecessary and that the term should be reassigned to its original Read Code, then the DCF would look as follows as Date2:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE

Term1	RC1	RC1	S	Date2
Term1	RC2	RC1	S	Date2

Note that Term1 attached to RC1 is now mapped back on to RC1. This is because Term1 will be attached to the analysis code RC2 if the DCF was previously applied and now needs mapping back on to RC1.

The lynchpin of this mechanism is the preserved selected Read Code in the clinical record, which can be revisited any number of times if new mappings become available. Re-allocation will not work if users replace the selected code with newly suggested codes. System developers should beware that Read_code_prev = Read_code_now does not mean 'No change'.

10.3 Redundancy followed by Improper Synonym

If a concept RC1 is made redundant to RC2 (Term1 moved from RC1 to RC2) at Date1, then the following entry is made in the DCF:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	R	Date1

If subsequently at Date2, Term1 is moved from RC2 to RC3 due to improper synonym, the DCF will be as follows.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	R	Date1
Term1	RC1	RC3	S	Date2
Term1	RC2	RC3	S	Date2

Note that the re-allocation due to redundancy is left in the table. This addresses two problems.

Firstly, some system designers may only wish to re-allocate Redundant codes (although this conservative stance is not advised).

Secondly, the following scenario may occur: a user makes an initial change (e.g. to correct for a Read Code being made redundant) and then makes a subsequent change (e.g. resolving an improper synonym). If the second change is later found to be inappropriate, the user may recover to the current and not the redundant code.

The same approach holds for re-allocation due to discovered redundancy followed by re-allocation to resolve ambiguity.

10.4 Term Made Obsolete followed by Concept Made Redundant

Initially Term1, which is attached to RC1, is made obsolete. The change is recorded in the DCF as the description is deleted from the Description file in the new release. But the obsolete term is simply 're-allocated' to its old code.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC1	O	Date1

Subsequently, the concept RC1 is made redundant to RC2. The original entry is deleted, and the redundant change is inserted in the DCF.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	R	Date2

Thus the Flag 'O' describes the possible actions to be taken in response to the re-allocation, and cannot be used as a resource to find all obsolete terms. A list of obsolete terms can be obtained by looking at the term status flag in the Term file or by comparing the Description file (containing all and only current terms) with the Terms file (containing all current and obsolete terms).

Once a term is made obsolete it may later be re-allocated to a new Read Code if the sense of the term is clear. For example, in the case of a spelling mistake, the term will shadow any future placements of its correct form. If, at Date3, an obsolete term is considered to be an improper synonym to RC2, and is re-allocated to RC3, the DCF will appear as follows:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	R	Date2
Term1	RC1	RC3	S	Date3

10.5 Redundancy (or other change) followed by Term Made Obsolete

If a concept RC1 is made redundant to RC2 (Term1 moved from RC1 to RC2) at Date1, then the following entry is made in the DCF:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	R	Date1

If Term1 is then made obsolete at Date2, no change is made to the DCF, i.e. the original record is left in place:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	R	Date1

Users and system designers will have already corrected for this, if they are actively using the DCF, so no further action is required. The unchanged date indicates this. In general, obsolete term changes are only flagged in the file if no other change accompanies making the term obsolete.

10.6 Ambiguity followed by Discovered Redundancy

If Term1 (a synonym in this case) attached to RC1 is found to be ambiguous, it may be assigned to RC2 or RC3, depending on the context of the patient case.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC3	A	Date1

If subsequently RC1 is found to be redundant to RC4, then the DCF will be as follows:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC3	A	Date1
Term1	RC1	RC4	R	Date2

Therefore, if ambiguity was not corrected in some or all cases after Date1, the redundant RC1 may still be replaced by the analysis code RC4, and the ambiguity corrected later. This works because even after automatically inserting the code RC4 in the analysis field, RC1 is still available in the selected field for matching onto the DCF. If the ambiguity had already been corrected, then there will be no instances of clinical records that need updating at Date2. To recognise this, a system must store the fact that the ambiguity has been resolved against each record containing the selected code RC1.

If, instead of RC1 being redundant to RC4, it was found that RC3 is redundant to RC5, then the DCF would be as follows at Date2:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC5	A	Date2
Term1	RC3	RC5	R	Date2

This last case highlights a way in which the flagging can be inefficient. Users who have resolved the ambiguity between RC2 and RC3 are now asked to distinguish between RC2 and RC5,

when all that is really needed is an automatic updating of the discovered redundancy. This is a side effect of DCF that is a 'snapshot' rather than a transaction file. However, the chosen format is considered more robust and the combinations that give rise to this inefficiency should not occur frequently.

10.7 Ambiguous Synonym

In some instances a synonym is found to be ambiguous, but the preferred term is unambiguous. As the preferred term defines the concept in CTV3, the original concept will remain valid. However, a new concept with another unambiguous preferred term will be added, and the user offered the chance of mapping to this alternative or of staying with the original choice of Read Code:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC1	A	Date1
Term1	RC1	RC2	A	Date1

10.8 Ambiguity followed by an Improper Synonym

Initially, Term1 (a synonym in this case) attached to RC1 is found to be ambiguous, and may be reallocated to RC2 or RC3, depending on the context of the case.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC3	A	Date1

Subsequently, Term1 is found to be an improper synonym to RC3, and is re-allocated to RC4.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC4	A	Date2
Term1	RC3	RC4	S	Date2
Term1	RC2	RC2	A	Date2
Term1	RC2	RC4	A	Date2

Two points arise in this case.

Firstly the flag on the new re-allocation from RC1 to RC4 is 'A' (rather than 'S'), as ambiguity requires more user input than resolving improper synonyms. This highlights that the flags are indicators of action to be taken.

Secondly, the user is also asked to check Term1/RC2 allocations as the improper synonym of RC4 may have wrongly driven a user to select RC2.

10.9 Improper Synonym followed by Ambiguity

Initially, Term1 is found to be an improper synonym of RC1 and re-allocated to RC2.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	S	Date1

Subsequently, Term1 is found to be ambiguous, and might go to RC2 or RC3.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date2
Term1	RC1	RC3	A	Date2
Term1	RC2	RC2	A	Date2
Term1	RC2	RC3	A	Date2

Once more, the A flag dominates over the S flag. Even if users have remapped to RC2 previously, they should reconsider this in the light of the new code RC3.

10.10 More Complex Combinations

It is possible that three or more changes will be made to a Read Code and term combination. The guidelines above will still apply, as in the following example.

If Term1 (a synonym in this case) which is attached to RC1 is found to be ambiguous, it may be reassigned to RC2 or RC3, depending on the context of the case.

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC3	A	Date1

If subsequently it was found that RC3 is redundant to RC5, then the DCF would be as follows at Date2:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC5	A	Date2

Term1	RC3	RC5	R	Date2

If then the combination of Term1 and RC5 was found to be an improper synonym, and Term1 should instead be mapped to RC6, the entry in the DCF would be as follows:

V3_TERM_ID	READ_CODE_PREV	READ_CODE_NOW	MAP_STATUS	RELEASE
Term1	RC1	RC2	A	Date1
Term1	RC1	RC6	A	Date3
Term1	RC3	RC5	R	Date2
Term1	RC3	RC6	S	Date3
Term1	RC5	RC6	S	Date3

Therefore redundant Read Codes which are superseded (e.g. RC3) are only stored as a 'fall-back' option in the DCF against the Read Code they are redundant to (RC5). Otherwise, only the best mappings are retained.

11. Where should users go for further advice?

Further implementation guidance notes are provided in Appendix 2. System developers are strongly advised to discuss migration and updating issues with the UK Terminology Centre, at the address given at the beginning of this document. Clinical users should first discuss their updating needs with their system suppliers/developers.

Appendix A

A List of Allowed Changes to Terms

Currently, the UK Terminology Centre allows minor changes to terms. This is occurring while a standard form is found for terms across the thesaurus e.g. *'gall bladder'* or *'gallbladder'*; *'Wilm's'* or *'Wilms'*.

A.1 Permitted Changes to Terms (in the short term)

- Adding or removing spaces
- Adding or removing hyphens
- Adding or removing apostrophes
- Changing capitalisation
- Correcting spelling mistakes

A.2 Changes not Permitted (example)

- Changing singular to plural or vice-versa
- Deleting a word (other than when adding or removing spaces, adding or removing hyphens)
- Adding new words (other than when adding or removing spaces, adding or removing hyphens)
- Changing a word to another word

Appendix B

Further Implementation Notes

There are many possible ways of implementing a patient record structure that supports an analysis field. The following examples are given as a brief guide. They are largely based on discussion rather than on existing implementations of medical records. Note that only core terms (and not qualifiers) are considered. Although the issue of qualifiers will be dealt with in a separate template change file document, the principles of clinical record construction are expected to be the same. The user is also directed to Section 9, where there is a discussion of potential flags against analysis Read Codes.

B.1. Analysis Field in Patient Notes File

This is a literal rendition of the advice given in this document and requires minimal change to the patient notes file.

Notes table
Event identifier
Selected Read Code
Selected term identifier
Selected term rubric
Date of recording
Analysis code
Date of recording analysis code
Potential ambiguity flag (see 11.3)
Auto flag (see 11.3)
(Other Flags, dates, audit trail info.)

However, there are disadvantages to this design:

- There is a serious issue as to whether it is wise from a medico-legal perspective to allow users to alter a field in a record once it has been written (even if the other archived fields for selected Read Code, term identifier and term cannot be edited).

- Compared with the Events log file (see below), it is incomplete, only recording the latest change. Subsequent changes would overwrite previous changes.

- Developers using this style of table will need to keep a separate audit trail of all user deletions/changes of events, and of all changes to the analysis code.

B.2. Parallel Analysis Table

By separating the analysis field out into a separate file, two problems are circumvented. The user never changes the original record when updating an analysis field. Also, the result is a relatively small sized analysis table which may be fast to search.

Notes table
Event identifier
Selected Read Code
Selected term identifier
Selected term rubric
Date of recording
Potential ambiguity flag (see 11.3)
(Other Flags, dates, audit trail info.)

Analysis table
Event identifier
Analysis Read Code
Date of recording
Auto flag (see 12.3)
(Other flags, dates, audit trail info.)

Some of the disadvantages of the single notes file method still hold:

- Compared with the Events log file (see below), it is incomplete, only recording the latest change. Older changes will be overwritten.
- Developers using this style of table will also need to keep a separate audit trail of all user deletions/changes of events, and of all changes to the analysis code.
- Denormalising (duplicating) the information will require extra storage space.

B.3. Addition and Deletion Log

In this type of clinical record file structure, all changes to the patient record are recorded in the same table as the original user's selection (the events table).

The core idea is that all records (user selections, automatic selections, user deletions) have the same structure. If a user 'deletes' an entry from the patient database, 'overwriting' it with a new entry, then the original record remains untouched, but a new record is inserted and flagged as *Current* selected. The

deleted record would be marked as *Previous* selected. Likewise, analysis codes may be marked as *Current* or *Previous* analysis codes.

If a sequence number is also included, users may see the sequence of selections and 'deletions' of both user selected and mapped codes used for information retrieval.

Event table fields	Field description
Event identifier	The identifier for this event.
Read Code	The Read Code of the entry – this may be a selected Read Code or an analysis Read Code.
Term identifier	The term identifier of the selected entry
Term rubric	The rubric of the selected term
Selected code flag	Flag denoting whether the user selected the code from the clinical information system (*Selected* code) or in response to a suggestion in the Description Change file (*Analysis* code). S = Current Selected code O = Old Selected code A = Analysis code P = Previous analysis code
Ordering number	Number denoting the order of the entry. This allows the sequence of changes to be reconstructed.
Auto flag	Flag denoting whether the entry was manual (e.g. if user selected, if ambiguity resolved in response to DCF prompt) or automatically inserted by the clinical information system (e.g. Redundancy) in response to DCF prompt. This flag records the decision that a choice has been made by the user, even if the choice was to stay with the original (e.g. ambiguous) Read Code.
Potential ambiguity flag	When the DCF is applied, clinical records that contain terms that are ambiguous may be flagged. Then the user may respond at a suitable time.
Date record created	The date the Read Code, etc. details were inserted into the record either automatically or by user selection.
Date of DCF used	There is no automatic way of knowing from the date of an entry which DCF was used. For example, the data may have come from another system via an electronic message in which an earlier version of the DCF had been used.
Other flags, dates, etc.	

The term rubric and term identifier are duplicated (undesirable redundancy) in this style of record when changes are inserted as a result of DCF prompting. However, the term may be different if a new record is inserted as a result of

the clinical record entry being 'overwritten' by the user.

The fields *'Date of DCF used'* and '*Potential ambiguity'* are only needed when a change is made due to a DCF prompt (undesirable redundancy).

Some system developers may wish to split the table into two tables with essentially the same structure. The first might contain user selections, the second automatic selections. This would solve the redundancy issues discussed above. System developers will need to consider the needs of reporting and analysis before deciding on whether or how the table is to be split.

This type of structure records all changes. Users can backtrack to the previous state of the record at any time.

B.4. Transaction Table

An alternative to the events log is a log of all changes made to the events file. The event table is maintained as an up-to-date file containing the latest information whether as a result of automatic or user selections. A separate file contains a list of all previous records, or of instructions that led to records being changed. This combination would contain precisely the same information as the Event log.

System developers and users may perceive this style as 'changing' the patient record and prefer the event log file. The style seems more error-prone.

B.5. Virtual Implementation of the Analysis Code

Another possibility is to avoid implementing the analysis code in the patient code. Instead the DCF is consulted when running a query. Users are asked to resolve ambiguity at this point.

For example, if the query is for all patients who have had the Read Code for a '*Myocardial infarct'*, then the user might be asked whether it is also intended that patients with '*Cardiac rupture'* are retrieved. If the user excludes these, then this fact can be added to the query.

- While the user's reply is stored, it is only a local decision to that query. If a later query for '*Ischaemic heart disease'* is run, the fact that the user is again asked whether or not '*Cardiac rupture'* needs to be excluded. This may be considered accurate, but may be relatively inefficient compared with deciding once and for all that '*Cardiac rupture'* is different from a '*Myocardial infarct'*.

- The algorithm must be carefully constructed. The search concept and all of its children (recursively down the hierarchy) should be checked to see if any of them match:

1) A description (*Read_Code_Prev* and *term_id*) that has changed. The query should be amended so that patient records matching *Read_Code_Prev* and *Term_id* are not recovered.

2) A new description (*Read_Code_now* and *term_id*), previously associated with a Read Code (*Read_Code_prev*) which is still found in patient records that would not ordinarily be recovered in the search. In this case *Read_Code_prev* must be added to the query if it is to be complete.

Note that searches must therefore be sensitive to the term identifier as well as the Read Code.

- Some decision support queries need to be made without constant user interruption, for example, detecting drug interactions, potential contraindications etc.

B.6. Text-based Systems (e.g. XML)

Although the paradigm is different, the general design and implementation principles still hold in SGML and XML-based systems. The selected Read Code might be tagged along with the term identifier and the term, and attached to the relevant body of text in the document. A further tag might identify an analysis code. However the analysis Read Code may have a virtual solution, implemented by mapping from the query through the DCF to the tags in documents constituting the patient record.

Appendix C

List of Major Changes to the Document

C.1. List of Changes from Version 1.1 of the Document

Following comment from many sources, an analysis of possible scenarios and testing at a live site, the following major changes have been made:

- The title of the file has been changed from Term Re-allocation File to Description Change file.
- Concept re-allocations has been separated from all other kinds of change in the release files to avoid any potential mishap.
- It is suggested that users change the 'selected' Read Code in the case of concept reallocation (only) to safeguard against possible confusion over the meaning of a Read Code. The originally selected Read Codes should still be archived for medico-legal purposes. This will happen automatically if system developers have implemented a transaction log for updates to clinical records.
- The mapping status flags have been changed, to better reflect the reason that the reallocation is being suggested, which is the most reliable indicator of the action required.
- Clearer guidelines are given on how the UK Terminology Centre will handle each type of re-allocation, along with combinations of re-allocation.
- Suggestions for implementation have been added.

C.2. List of Changes from Version 1.4 of the Document

- Principle of exchange of information modified; other principles clarified.
- No other major changes to the principles, file structure or examples.
- Clarification of the role of implementation advice.
- Addition of advice about virtual implementation of the analysis field and text based patient records.