Medical Coding in Clinical Trials

Deven Babre

ABSTRACT

Data generated in all clinical trial are recorded on the data collection instrument Case report Form / Electronic Case Report Form by investigators located at various sites in various countries. In multicentric clinical trials since different investigator or medically qualified experts are from different sites / centers recording the medical term(s) uniformly is a big challenge. Medical coders from clinical data management team process these terms and perform medical coding. Medical coding is performed to categorize the medical terms reported appropriately so that they can be analyzed/reviewed. This article describes process which is used for medical coding in clinical data management and two most commonly used medical dictionaries MedDRA and WHO-DDE in brief. It is expected to help medical coders to understand the process of medical coding in clinical data management. Few common issues which the medical coder faces while performing medical coding, are also highlighted.

Key words: Clinical Data Management, Medical Coding, MedDRA®, WHO-DDE, Verbatim Term, Medical Coding Dictionaries, Auto Coding, Manual Coding

DATA generated in all clinical trials is recorded on data collection instruments (DCIs) called as Case Record Forms / Report Forms (CRFs) in respect of paper based trials or as electronic Case Record/ Report form (eCRF) in respect of web based clinical trials. In these trials information on Adverse Events (AE), Medical History (MH), and Concomitant Medications (CM) used in addition to the study medication are collected and recorded on relevant DCIs. In a multicentric clinical trial there are many trial sites, which involve different investigator with different ethnic backgrounds. It is anticipated that due to involvement of investigators and clinical research professionals from different countries/region there is a possibility of recording medical/scientific data in different fashion. All data generated in these trials are ultimately subjected to further analysis. It is very essential that this data gets interpreted uniformly in a standardized format. Hence medical coding is required by using standardize medical dictionaries. Data listed above like AEs, SAEs, MH , CM and any other category generally are coded. However coding AEs, SAEs and CM is mandate in any given clinical trial.

In this article the attempt is made

- to explain how medical coding is performed in clinical trial using Standardised medical coding dictionaries.
- to describe in brief two most extensively used dictionaries by most of the regulatory offices and majority of professionals in Pharmaceutical/CRO industry and
- to describe process of medical coding and also to share some of the common problems that a coder faces during medical coding.

There are several standardized medical coding dictionaries in the market; however five dictionaries listed below are used for coding:

1. **COSTART** - Coding Symbols for Thesaurus of Adverse Reaction Terms
2. **ICD9CM** - International Classification of Diseases 9 Revision Clinical Modification
3. **MedDRA** – Medical Dictionary for Regulatory Activities
4. **WHO-ART** – World Health Organisation Adverse Reactions Terminology
5. **WHO-DDE**– World Health Organisation Drug Dictionary Enhanced

Out of the above five, two widely used medical coding dictionaries used for coding medical terms generated in clinical trials are MedDRA and WHO-DDE. To maintain uniformity in reporting a term is next to impossible in any given clinical trial. However for a coder it is a challenging task to ensure that the term recorded/reported on data collection instrument (CRF/eCRF) is coded appropriately.

It is well known fact that these dictionaries come at a price and the organisations that perform medical coding activity should posses appropriate valid licences. There are separate licences issued to separate user groups for each of the dictionaries appropriately.

Processes in medical coding

Precoding Process:

Any medical coding dictionary and all subsequent

Deven Babre
Director, Data Management, PharmaNet Clinical Services Pvt Ltd.

revisions have to be correctly imported / loaded in the appropriate coding tool by the database programming team. In respect of Oracle Clinical (OC) the coding tool used is Thesaurus Management System (TMS). Once the dictionaries are imported / loaded in the appropriate tool the programming team checks if all the tables/records are correctly loaded in the tool. This process is performed only once for the given version of dictionary. After ensuring correct import tables and records in the tool by the programming team, the operational team performs user acceptance test (UAT) in which the members of operational team confirm that the dictionary which is loaded in the tool is giving the required output as expected. Once the operational team clears UAT, the selected dictionary is released for use for a particular project/study. If the same version of dictionary is to be used for any another project/study in future, then UAT needs to be repeated by members on the operational team assigned to the new project / Study.

Prior to assigning a dictionary to a project/study, it is essential to check the following points as a prerequisite:

- Latest validated version available in the coding tool at the start of the project
- Policy / requirement about use of the same version of dictionary available in the tool for entire project in spite of availability of the newer version
- upgraded version prospectively as and when available during the life of the project
- upgraded version retrospectively as and when available during the life cycle of a project.

Coding in Live Project

Ideally coding is performed on validated and cleaned data by data managers responsible for "Data Review and Discrepancy Management". The terms to be coded in any project get coded by a process known as "auto coding", the terms which fail to get "auto-coded" have to be coded "manually" by medical coder responsible for the project. Two processes of coding – auto and manual are described below in brief:

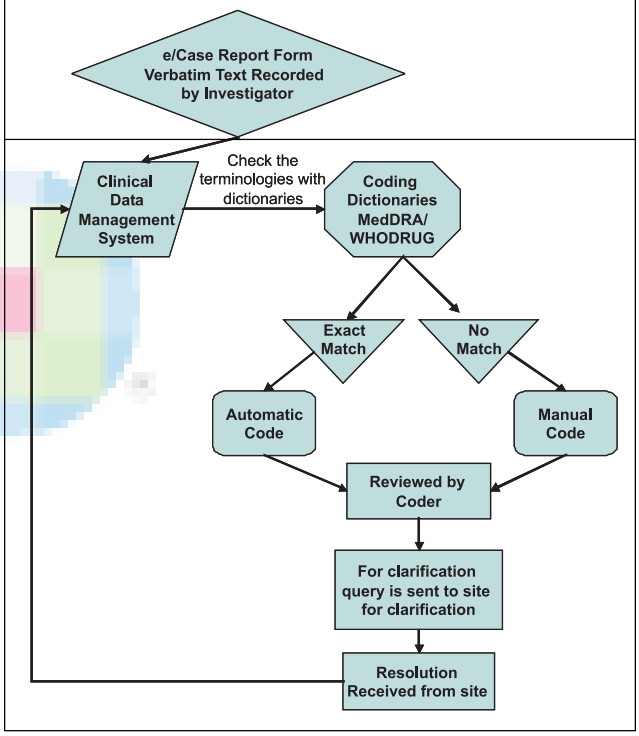
Auto Coding: The term recorded by the investigator on the data collection instrument gets coded automatically if it exactly matches with the appropriate term available in the medical dictionary.

Manual Coding: Auto coding fails in respect of terms which do not match with the appropriate level of hierarchy in the medical dictionary. All these terms are required to be manually coded by the medical coder assigned to the project. The medical coder will find the appropriate match for the term from among the terms within the assigned dictionary and will manually assign the code.

This does not mean that all terms reported and recorded on CRF / eCRF get coded without any issues. There are some terms which are unclear or for which it is not very easy for a coder to find matching term within the dictionary. The investigator may report multiple signs and symptoms. In such cases, the medical coder / medical coding team sends these terms to investigator / medically qualified experts for clarification/more information. It helps the medical coder to identify term(s) very close to such unclear or doubtful terms within the coding dictionary so that the term(s) get appropriately coded. Term(s) which get auto and manually coded are reviewed by the coding personnel. Unclear term(s) / term(s) with insufficient details are queried to site. Investigator must provide appropriate updates/details and send the signed resolution back to data management team. Based on investigator resolution, the data management team takes appropriatel action in database. The coder looks at the information/update and then codes the term appropriately.

The same process is depicted in a flow chart below

Fig 1 Medical Coding Process



Medical Dictionary for Regulatory Activities (MedDRA)

Medical Dictionary for Regulatory Activities (MedDRA®) is a medical coding dictionary developed by Maintenance and Support Services Organisation (MSSO). MedDRA® is supported by International Conference on Harmonisation (ICH) on Technical Requirements for Registration of Pharmaceuticals for Human use. Prior to development of MedDRA, there was no internationally accepted medical terminology for biopharmaceutical regulatory purposes.

MedDRA is used for coding

- medical terms generated during all phases of clinical trial, excluding animal toxicology,

- therapeutic indications which include signs, symptoms, diseases, diagnosis, or prophylaxis of disease, and modification of functions,
- coding names and quantitative results of investigations, surgical procedures and medical/social/family history.

MedDRA releases 2 versions in a year - one in March and the second in September. One can obtain access to the MedDRA terminology annually, by renewable subscription. Each subscription brings all MedDRA updates that incorporate approved changes and additions.

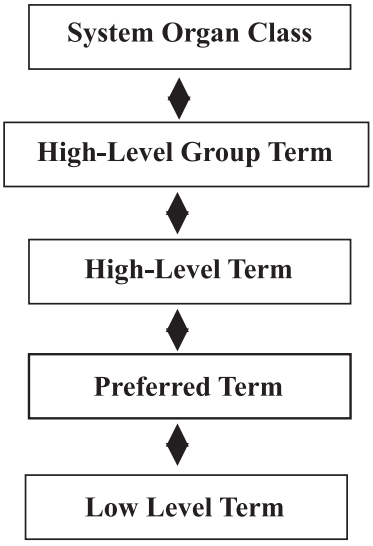
MedDRA has five hierarchical levels as listed below (Fig 2)

- Low Level Term (LLT)
- Preferred Tem (PT)
- High Level Term (HLT)
- High Level Group Tem (HLGT)
- System Organ Class (SOC)

Low Level Term (LLT) is the lowest level of the terminology. Each LLT is linked to only ONE PT. A PT distinctly describes a symptom, sign, disease, diagnosis, therapeutic indication, investigation, surgical, or medical procedure and medical, social, or family history characteristics[1].

High Level Term (HLT) is a superordinate descriptor for PTs linked to it. High Level Group Term (HLGT) is a superordinate descriptor for one or more HLTs related by anatomy, pathology, physiology, etiology, or function. System Organ class (SOC) is the highest level of hierarchy. The SOCs are grouped by etiology, manifestation site and purpose[1].

Fig 2 Structural Hierarchy of the MedDRA® Terminology [1]



Common problems faced by medical coding expert while coding:

- illegible verbatim term
- spelling errors
- use of abbreviations
- multiple signs and symptoms recorded as separate events which may lead to some diagnosis (for example: signs and symptoms recorded as running nose, cough and fever may lead to diagnosis of Pneumonia)
- multiple medical concepts recorded together. To code we need to split the terms.
- event is recorded without mentioning the site e.g. ulcer is recorded without additional information like moth ulcer, leg ulcer etc
- multiple medical concepts recorded which had surgical procedure and reason for injury. However the reason or cause or site of injury is not clarified.
- a medication term reported however allergy due to the medication or outcome of the allergy is not specified

Other commonly used dictionaries

World Health Organisation Drug Dictionary (WHODRUG): This is a dictionary maintained and updated by Uppsala Monitoring Centre (UMC). This dictionary is most comprehensive dictionary which has medicinal product information. It is used by drug regulatory authorities, various pharmaceutical companies and contract research organizations (CROs). The dictionary covers medicinal product names – proprietary and non- proprietary – from more than 90 countries. WHODRUG dictionary has undergone a lot of development. Currently we have of three dictionary types[2]

1. WHO Drug Dictionary (WHO-DD)
2. WHO Drug Dictionary Enhanced (WHO-DD Enhanced)
3. WHO Herbal Dictionary (WHO-HD)

The WHO DD and WHO DD Enhanced mainly contain information about conventional medicinal products, but a number of other product types listed below:

- Medicinal product
- Herbal remedy
- Vaccine
- Dietary supplement
- Radio-pharmaceutical
- Blood product
- Diagnostic agent
- Homeopatic remedy

The WHO Herbal Dictionary contains almost all herbal entries that have been entered into WHO Drug Dictionary over the years. From 2005 all herbals will be included exclusively in the WHO Herbal Dictionary. The WHO Herbal Dictionary is classified with the Herbal Anatomical Therapeutic Chemical (HATC) classification[2].

The WHO Drug Dictionaries contain information about the Medicinal Products. This information is used to identify a term (medicinal product) closely matching with the term reported on DCI.

ATC classification is integral part of the dictionary. This is used to classify the medicinal product to the main therapeutic use of the active ingredient/s

LEVEL 1: Anatomical main group

LEVEL 2: Therapeutic subgroup

LEVEL 3: Pharmacological subgroups
LEVEL 4: Chemical subgroups
LEVEL 5: Chemical substance

As per the above system the ATC classification for medicinal product Metformin will be as follows:

A	Alimentary tract and metabolism (1st level, anatomical main group)
A10	Drugs used in diabetes (2nd level, therapeutic subgroup)
A10B	Blood glucose lowering drugs, excl. insulins (3rd level, pharmacological subgroup)
A10BA	Biguanides (4th level, chemical subgroup)
A10BA02	Metformin (5th level, chemical substance)

Common problems faced by medical coding expert while coding medicinal products:

- illegible verbatim term
- spelling errors
- use of abbreviations
- indication prescribed for the medicinal product is not an approved indication mentioned on prescribing information
- local brand available in market and generic/active ingredient is not known.
- multiple medications recorded together. To code we need to split the terms.

Acknowledgements

I wish to thank Dr. Suresh Bowalekar, Ph. D., C. Stat. (UK), C. Sci (UK); Managing Director, PharmaNet Clinical Services, Pvt Ltd for all his help and guidance.

References:

1. MedDRA Introductory Guide Version 12.1, September 2009, MSSO-DI-6003-12.1.0
2. WHO Drug Dictionary Sample Getting Started © UMC Products & Services, 2006

Do not become archivists of facts. Try to penetrate to the secret of their occurrence, persistently search for the laws which govern them.

— Ivan Pavlov