G. Jaya Nair

*Independent Consultant,
Clinical Data Management,
New Jersey, USA.*

Address for correspondence:
Dr. Jaya Nair G,
35 Eastwood Drive,
East Windsor, NJ 08520, USA.
E-mail: jayagnair@gmail.com

Ensuring quality in the coding process: A key differentiator for the accurate interpretation of safety data

Abstract

Medical coding and dictionaries for clinical trials have seen a wave of change over the past decade where emphasis on more standardized tools for coding and reporting clinical data has taken precedence. Coding personifies the backbone of clinical reporting as safety data reports primarily depend on the coded data. Hence, maintaining an optimum quality of coding is quintessential to the accurate analysis and interpretation of critical clinical data. The perception that medical coding is merely a process of assigning numeric/alphanumeric codes to clinical data needs to be revisited. The significance of quality coding and its impact on clinical reporting has been highlighted in this article.

Key words: Clinical trials, dictionaries, dictionary hierarchies, MedDRA, medical coding, quality, safety signals, World Health Organization Drug

INTRODUCTION

Globalized clinical trials have augmented the need to systematize data collected over various geographies with native languages and local drugs for consistent reporting. Quality of clinical data and reports is the foundation for the precise reporting of clinical trial information.

Coding of clinical data is the conduit to consistent and standardized retrieval of medical conditions/information. It is the source to identify the frequency of adverse drug reactions/adverse events (AEs) and provides common denominator for comparison across different trials. The coded data forms one of the key components of the tables and reports required for regulatory submission. Hence, the understanding that the function of coding is not trivial is essential and the need to maintain the high quality in the delivery of coded data should not be underestimated.

Data collected as AEs/serious AEs/medical histories/

surgical procedures etc., and concomitant medications are extremely critical domains that form the basis of safety and efficacy of the investigational drug or for post-marketing. In global trials, the data on these pages may be captured in different ways and the drugs used in various geographies may be named differently as well. Hence coding of this data into medical dictionaries such as Medical Dictionary for Regulatory Activities (MedDRA) and WHODrug provide a platform to represent different terminologies under a standard-consistent umbrella to detect any safety signals that may be associated with the drug.

Dictionaries are repositories of medical terminology and the growing emphasis of using standard dictionaries such as MedDRA and WHODrug mandates the need for the coder to have a thorough understanding of them. Understanding the hierarchies and classifications is the pre-requisite for accurate coding. Good to support this with an example. A lack of understanding of the dictionaries may pose a serious threat to the quality of coding. Key skills and capabilities are required to perform a task as critical as coding. As mentioned in the "data retrieval and presentation (DRP): Points to consider" document, unless users achieve consistency in how they assign terms to verbatim reports of symptoms, signs, diseases, etc., and in methods for data retrieval and evaluation, use of MedDRA cannot have the desired harmonizing effect in the exchange of coded data. Although a highly granular terminology

like MedDRA reduces the need for interpretation at data entry, it impacts the processes of data retrieval, sorting, and presentation necessary for the support of drug development, pharmacovigilance and risk management.[1]

MedDRA – the Medical Dictionary for Regulatory Activities – is a medical terminology used to classify the AE information associated with the use of biopharmaceuticals and other medical products (e.g., medical devices and vaccines). Coding these data to a standard set of MedDRA terms allows health authorities and the biopharmaceutical industry to more readily exchange and analyze data related to the safe use of medical products.[2] Add reference as this is taken as is from MSSO MedDRA site It is used to report AE data from clinical trials, and for post-marketing reports and pharmacovigilance. MedDRA is a pragmatic, clinically validated terminology that applies to all phases of drug development, and excluding animal toxicology and it also applies to the health effects and malfunction of medical devices.[3]

MedDRA hierarchy

The hierarchical structure of MedDRA facilitates data retrieval by providing grouping terms (High Level Terms [HLTs] and High Level Group Terms [HLGTs]) that aggregate the very specific terms used for coding into broader medical categories [Figure 1].[1]

MedDRA multiaxiality

Multi-axiality means that a preferred term (PT) may exist in more than one system organ class (SOC). This allows terms to be grouped in different, but medically appropriate, ways (e.g., by etiology or organ system). Each PT is assigned one primary SOC; all other SOC assignments for that PT are called "secondary." Having a single primary SOC prevents double counting of events when outputting data from all SOCs [Figure 2].[1]

WHO Drug Dictionary Enhanced (DDE) – WHODDE's hierarchical product coding system, its range of powerful analytical tools, and its extensive coverage make it a valued means of interpreting and reporting medicinal product information. Most importantly, WHODDE meets the expressed need for a consistent drug dictionary and exact terminology. With WHODDE, users are able to code concomitant medication, better analyze and understand the resulting data, and accelerate submissions to regulatory authorities. Drug safety surveillance is also enhanced. The quality and correctness of their data is enhanced, and the resulting benefits are seen throughout drug development and safety surveillance – both pre- and post-marketing. National regulatory authorities, which have a natural interest in the fast, safe and correct communication of clinical and drug safety data collected from trials and reports from around the world are also regular users of WHO DDE.[4]

The structure of WHO Drug provides a clear understanding of the drug.

The Anatomical Therapeutic Chemical (ATC) classification system, the active substances are divided into different groups according to the organ or system on which they act and their therapeutic, pharmacological and chemical properties. Drugs are classified in groups at five different levels. The drugs are divided into fourteen main groups (1st level), with pharmacological/therapeutic subgroups (2nd level). The 3rd and 4th levels are chemical/pharmacological/therapeutic subgroups and the 5th level is the chemical substance. The 2nd, 3rd, and 4th levels are often used to identify pharmacological subgroups when that is considered more appropriate than therapeutic or chemical subgroups.

The complete classification of metformin illustrates the structure of the code:
An 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
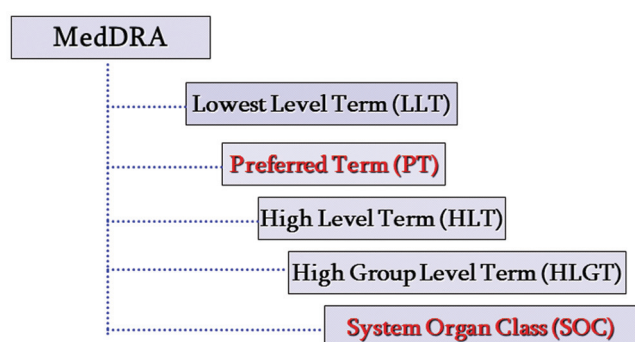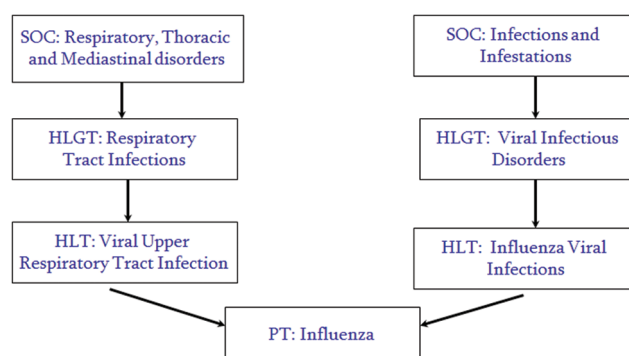(3rd level, pharmacological subgroup)

Figure 1: MedDRA hierarchy

Figure 2: MedDRA multiaxiality

A10BA Biguanides
(4th level, chemical subgroup)
A10BA02 metformin
(5th level, chemical substance).

Thus, in the ATC system all plain metformin preparations are given the code A10BA02.[5]

Thus, MedDRA is used to report AE data from clinical trials as well as post-marketing and pharmacovigilance data. WHODDE is a drug dictionary, which is useful when tabulating medication and its usage as the dictionary has the structure that classifies the same medication, often known by different names, into a single name.

Other dictionaries like international classification of diseases, Coding Symbols for Thesaurus of Adverse Reaction Terms and World Health Organization Adverse Reactions Terminology were also used for coding; however, due to the need for more precise and standardized reporting and data retrieval MedDRA dictionary is being used for coding in clinical trials.

Coding revolves around some of the main principles:
• Clearly defined and documented coding process.
• Quality assurance.
• Quality of source data.
• Level of term selection.
• No addition or subtraction of information.

Clearly defined and documented coding process
The process of coding should be clearly defined and followed to maintain consistency. The points to consider document serves as a good basis to create trial specific coding conventions. The coders should be completely aware and knowledgeable about the conventions and process of coding. To promote consistency, organizations should document their term selection methods and quality assurance procedures in coding guidelines consistent with MTS (MedDRA Term Selection): Points To Consider (PTC) document.[6]

Similarly, the process of up-versioning and dictionary maintenance should be documented and followed. Each organization should have a versioning strategy that should be documented. The versioning strategy may differ between safety databases and clinical trial databases. For example, there may be no need to update clinical trial data from older trials if the data are not presently used or will not be used in the future. On the other hand, post marketing safety data may need to be reported in the current (or near-current) version of MedDRA, and version update recommendations would then apply.[6] Lack of standard processes, work-flows, and evaluation measures impact the quality of coding. Organizations are encouraged to document their data retrieval and output strategies, methods and quality assurance procedures in organization-specific guidelines, which should be consistent with this DRP: PTC document.[1]

Quality assurance
The quality of coded terms should be assessed either by quality review (QC – Quality Control process) or by medical review to ensure accuracy. Individual coder accuracy and consistency within the team of coders should be evaluated critically and a process to maintain this quality should be documented. Standard monthly metrics and evaluation tools must be maintained for each coder in the team as part of performance assessment; this would help in maintaining a check on quality and individual development. This in turn will also help for identification of specific training needs and challenges faced by the team that may affect quality and appropriate remedial actions can be taken to avoid replication of errors. High quality data output occurs when the quality of the information originally reported is maintained with consistent and appropriate term selection. Organizations should ensure continuous oversight of data quality.[1] The quality of the reports is critical for the appropriate evaluation of the relationship between the product and AEs.[7] Creation of standard work-flows, defining benchmarks for quality as well as continuous improvement will contribute to the improvement of the quality of coded data.

The quality of source data
One of the challenges posed is that the quality of the source data as sometimes ambiguous; extraneous or confusing information may be collected. A meaning that is clear to the investigator at the point of data entry in the context of the study may be unclear for universal coding. Unacceptable abbreviations, multiple interpretations/not enough information for classification are common challenges that are observed, undocumented/conflicting drug names, the same trade name being used for different products, spelling variations especially for drug names that can more easily be misinterpreted from actual drug name, all of these may present as additional factors that affect the quality of coding. The coder should be trained well enough to query such information accurately. The collection of clean data at the start can be promoted through the careful design of data collection forms, and the training of individuals on data collection and follow-up (e.g., investigators, drug sales representatives).[6]

The level of term selection
Dictionaries have different hierarchies and axialities (can you elaborate on this in brief). Please modify that would categorize the coded term into specific organ systems or ATC class.

Hence, the coder should be aware of the various levels of term classification and code the term appropriately. If the term could be coded to several different hierarchies and then the term should be queried as needed to request for further information to facilitate the exact classification of the term.

For example, the degree of specificity of some MedDRA Low Level Terms (LLTs) may be challenging for term selection.

Here are some tips for specific instances:
A single letter difference in a reported verbatim text can impact the meaning of the word and consequently the term selection [Table 1].

The term selection principles used:
- Selecting more than one term when coding a medical condition increases the count of terms
- Selecting a diagnosis term only (and not terms for signs and symptoms) reduces the count of terms.
- The AE profile resulting when both diagnosis and signs/symptoms terms are coded may appear different than when the diagnosis alone is coded. Always consider the organization's coding conventions when using or comparing data from other databases (e.g., co-developing or co-marketing partners, regulatory authorities).[1]

No addition or subtraction of information
At no point of time should a coder add or subtract information and code terms based on his her perception without having queried for the term and received confirmation from the investigator. There should not be any manipulation of data without an appropriate query response. If a diagnosis is reported with characteristic signs and symptoms, then the verbatim term should be coded to the diagnosis. When selecting terms, no reported information should be excluded from the term selection process; similarly, one should not add information by selecting a term for a diagnosis if only the signs or symptoms are reported.[2]

As coded data has a significant role to play in clinical reporting, the necessity of building certain key skills and capabilities in coders is essential. The coders should have an understanding of medical and scientific terminology.

They must also have the understanding of regulatory requirements and quality control and must have good analytical skills and attention to detail. The coder must be well-versed with ICH GCP (International Conference on Harmonisation - Good Clinical Practice) and the points to remember document to ensure accurate coding. The coders should also refer to different medical terminology books such as Dorland's Dictionary, Steadman's Dictionary etc., Not sure what you wish to express what you recommend the coder to refer to please specify. The coder must also have good communication skills in order to query for ambiguous or incomplete information. Intensive training and quality monitoring play a crucial role in developing the best and acceptable quality of coding.

The MedDRA coded data helps to create the Standard tables using the primary SOC view including HLGTs, HLTs and PTs (clinical trials and postmarketing data) and for cumulative summaries (postmarketing data). Line listings (both clinical and postmarketing data) can also be displayed by primary SOC and PT. The WHODrug coded data helps to detect possible drug interactions, protocol violations, prohibited medications, and serves as an effective tool for screening of "double medication" and "pseudo double medication." Coded data also helps for the early detection of safety signals and to observe data trends.

Data retrieval is performed for the summary and analysis of clinical trial data, pharmacovigilance, medical information questions and for a number of other purposes.[1] This data assists to identify safety signals that may warrant further investigation. Such data would include:
- New unlabeled AEs, especially if serious;
- An apparent increase in the severity of a labeled event;
- Occurrence of serious events thought to be extremely rare in the general population;
- New product-product, product-device, product-food, or product-dietary supplement interactions;
- Identification of a previously unrecognized at-risk population (e.g., populations with specific racial or genetic predispositions or co-morbidities);
- Confusion about a product's name, labeling, packaging, or use;
- Concerns arising from the way a product is used (e.g., AEs seen at higher than labeled doses or in populations not recommended for treatment);
- Concerns arising from potential inadequacies of a currently implemented risk minimization action plan;
- Other concerns identified by the sponsor or FDA.[7]

Reports retrieved also help to identify new indications for which the drug may seem useful leading to new scope for a clinical trial. Similarly, data procured can then easily be translated to create package inserts for each drug.

Table 1: Reported verbatim term vs LLT/PT

Example reported	LLT selected
Lip sore	Lip sore (PT lip pain)
Lip sores	Sores lip (PT cheilitis)
Sore gums	Sore gums (PT gingival pain)
Sores gum	Sores gum (PT gingivitis)[2]

LLT: Low level term, PT: Preferred term

CONCLUSION

Coding; thus, is an essential to improve consistency in comparing "safety signals" and aggregated clinical data. Hence, quality control should be applied to each stage of data handling to ensure that all data are reliable and have been processed correctly as this would directly impact safety information. Training should be provided to all coders in order to achieve the optimum level of coding and to ensure that all the parameters for quality are achieved. Coding should not be treated as a clerical function, but as one of the most important function in clinical research. Try adding what is your inference and conclusion/recommendations.

REFERENCES

1. MedDRA® data retrieval and presentation: Points to consider (ICH– endorsed guide for MedDRA users on data output – Release 3.3 Based on MedDRA version 15.0-1April 2012). (Reference page numbers 1, 8, 16, 5, 7, 14).
2. MedDRA MSSO. Available from: http://www.meddramsso.com/.
3. Available from: http://www.meddramsso.com/public_about_meddra.asp.
4. Available from: http://www.umc-products.com/DynPage.aspx?id=73588 and mn1=1107 and mn2=1139-UMC.
5. WHO Collaborating Centre for Drug Statistics Methodology, Guidelines for ATC Classification and DDD Assignment 2011. 14th ed. Oslo; 2010. (Reference page 15).
6. MedDRA® Term Selection: Points To Consider (ICH-Endorsed Guide for MedDRA Users Release 4.3 Based on MedDRA Version 15.0-1 April 2012) (Reference page numbers 2, 43, 3, 6).
7. Guidance for Industry-Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment (March 2005 Clinical Medical) (Reference pages 4 and 10).