PhUSE 2015

Paper IS04

Lot more than a patient's data

Satyapal Gudla, Novartis Healthcare Pvt. Ltd., Hyderabad, India
Nigel Montgomery, Novartis Pharma AG, Basel, Switzerland

ABSTRACT

In clinical trials, it's not just the patient's data that SAS programmers deal with. There are several other associated data/ information that are being used to analyze the clinical trial data. Patient's data are basically obtained from CRFs (Case Report Forms) which are processed and reported in a way that helps derive meaningful conclusions about the trial. This process of transforming patient's data into analysis often requires supplementary data/ information from different sources like dictionaries, reference or grade ranges, safety data specific to drug/ disease, any other information from past trial conduct experiences and/or literature etc. We refer to this supplementary data as non-CRF data.

This paper will give an overview of such non-CRF data, covering its importance in carrying out analysis with the examples demonstrating the basic usage. This will help new programmers in the industry to familiarize with the different set of data we work with/ refer to, while performing the analysis.

INTRODUCTION

The purpose of this paper is to let new statistical analysis programmers in the pharmaceutical industry understand that it's not just the data collected from patients we work with, but there are several other references, in this highly research oriented industry, that are used to derive clinically meaningful and scientifically accepted inferences about the trial.

Disclaimer:

The paper focuses on demonstrating the basic purpose of such references and how that contributes to the trial analysis with easy to understand examples for industry starters, at a quick glance, without giving too many details. This paper does not cover any programming aspects in working with these references. The content mentioned is solely from the authors' experience in these areas.

Assumptions:

The reader understands the clinical trial data, for example adverse events, risks, concomitant medications, laboratory and vital signs.

DICTIONARIES: MEDDRA AND WHODRUG

You would probably hear about dictionaries when you work with adverse events, medical history and concomitant medications data. In practice, the investigators across the sites would always try to describe patients' symptoms in as much detail as possible in the free text on the CRF. For example: for one patient it is 'BACK PAIN', for another 'BACK PAIN LEFT SIDE', or 'BACKACHE' or 'BACK INJURY' and for another patient it could be more descriptive as 'BACK PAIN RELATED TO MUSCLE STRAIN'. For the analysis, we cannot really use these reported terms directly as the list would be enormous and no good inferences could be made from it. Thus, there is a need for a clinically meaningful grouping or classification of these AE terms and this can be achieved using a dictionary named MedDRA (Medical Dictionary for Regulatory Activities). Similarly, for medications data, it is very common that one drug/ medication can have several brand/ trade names in the market. For example: 'ASPIRIN', 'DISPRIN', 'COLDRIN', 'CETIRIZINE' etc. which again demands a grouping for better analysis of medications data and WHODrug (World Health Organization Drug) dictionary helps us with this classification of drugs. Both these dictionaries are clinically meaningful, validated and accepted across the industry.

Below tables illustrate the basic structure of MedDRA and WHODrug dictionaries.

Table 1.1: Classification of adverse events using MedDRA dictionary

Body System Organ Class (SOC)	MUSCULOSKELETAL AND CONNECTIVE TISSUE PAIN AND DISCOMFORT	INJURY, POISONING AND PROCEDURAL COMPLICATIONS	MUSCULOSKELETAL AND CONNECTIVE TISSUE PAIN AND DISCOMFORT
High Level Group Term (HLGT)	MUSCULOSKELETAL AND CONNECTIVE TISSUE PAIN AND DISCOMFORT	INJURIES NEC	MUSCULOSKELETAL AND CONNECTIVE TISSUE PAIN AND DISCOMFORT
High Level Term (HLT)	MUSCULOSKELETAL AND CONNECTIVE TISSUE PAIN AND DISCOMFORT	SITE SPECIFIC INJURIES NEC	MUSCULOSKELETAL AND CONNECTIVE TISSUE PAIN AND DISCOMFORT
Preferred Term (PT)	BACK PAIN	BACK INJURY	MUSCLE STRAIN
Lowest Level Term (LLT)	BACKACHE	BACK INJURY	MUSCLE STRAIN
Reported AE Term	BACKACHE	Back Injury	Back pain related to muscle strain

In the above table, the verbatim AE terms in CRF, "BACKACHE" and "Back pain related to muscle strain" are classified to same System Organ Class.

Points to note:

- The hierarchy of the MedDRA dictionary is organized, from very specific to very general i.e. Lowest level term (LLT) to Preferred term (PT) to High-level term (HLT) to High-Level Group term (HLGT) and then to the System organ class (SOC).
- Each of the MedDRA terms, i.e. LLTs through SOCs are associated with the numeric codes.
- The CRF reported AE term is mapped to LLT code of the lowest level term that matches/ nearly matches by the medical coding specialists. This then serves as a basis to retrieve the remaining associated MedDRA hierarchy for that event.
- Each LLT is linked to only one PT. One PT can be linked to multiple SOCs via different HLT, HLGT paths. This feature of MedDRA dictionary is called "Multi-axiality", that allows a term to be represented in more than one SOC i.e. Primary SOC and one or more secondary SOCs.

  For example: Preferred Term Influenza represents an important respiratory tract problem as well as an infection (as shown below).

Table 1.2: Multi-Axial feature of MedDRA dictionary

Body System Organ Class (SOC)	INFECTIONS AND INFESTATIONS	RESPIRATORY, THORACIC AND MEDIASTINAL DISORDERS
High Level Group Term (HLGT)	VIRAL INFECTIONS DISORDERS	RESPIRATORY TRACT INFECTIONS
High Level Term (HLT)	INFLUENZA VIRAL INFECTIONS	VIRAL UPPER RESPIRATORY TRACT INFECTIONS
Preferred Term (PT)	INFLUENZA	

PT Influenza is primary to the SOC Infections and infestations, but this PT is also secondary to the SOC Respiratory, thoracic and mediastinal disorders.

Further details about the MedDRA can be found at link mentioned in references section.

Table 2.1: Classification of drugs/ medications using WHODrug

ATC Level 1 (Anatomical)	RESPIRATORY SYSTEM	RESPIRATORY SYSTEM
ATC Level 2 (Therapeutical)	ANTIHISTAMINES FOR SYSTEMIC USE	COUGH AND COLD PREPARATIONS
ATC Level 3 (Pharmacological)	ANTIHISTAMINES FOR SYSTEMIC USE	COUGH SUPPRESSANTS, EXCL. COMBINATIONS WITH EXPECT
ATC Level 4 (Chemical)	PIPERAZINE DERIVATIVES	OTHER COUGH SUPPRESSANTS
Preferred Term	CETIRIZINE	CLOFEDANOL HYDROCHLORIDE
Reported medication	Cetirizine	COLDRIN

In the above table, different drug/trade names "Cetirizine" and "COLDRIN" are grouped to RESPIRATORY SYSTEM at ATC classification of level 1.

Points to note:

- WHODrug dictionary is based on ATC classes (A–Anatomical, T-Therapeutical/ Pharmacological and C-Chemical).
- It classifies a drug according to the organ/ system the drug acts on and its therapeutic, pharmacological and chemical properties.
- One drug is mapped to one preferred term, which can be coded to several ATC classes depending upon its therapeutic application and there is no "Primary ATC code" (Please refer to example in below table 2.2). We report all of the ATC classes that medication is classified under.
- Drug code in medication data is linked with ATC code in WHODrug dictionary similar to MedDRA coding.

Table 2.2: More than one ATC classification per drug

ATC Level 1 (Anatomical)	ALIMENTARY TRACT AND METABOLISM	BLOOD AND BLOOD FORMING ORGANS	MUSCULO-SKELETAL SYSTEM	NERVOUS SYSTEM
ATC Level 2 (Therapeutical)	STOMATOLOGICAL PREPARATIONS	ANTITHROMBOTIC AGENTS	TOPICAL PRODUCTS FOR JOINT AND MUSCULAR PAIN	ANALGESICS
ATC Level 3 (Pharmacological)	STOMATOLOGICAL PREPARATIONS	ANTITHROMBOTIC AGENTS	TOPICAL PRODUCTS FOR JOINT AND MUSCULAR PAIN	OTHER ANALGESICS AND ANTIPYRETICS
ATC Level 4 (Chemical)	OTHER AGENTS FOR LOCAL ORAL TREATMENT	PLATELET AGGREGATION INHIBITORS EXCL. HEPARIN	PREPARATIONS WITH SALICYLIC ACID DERIVATIVES	SALICYLIC ACID AND DERIVATIVES
Preferred Term (PT)	ACETYLSALICYLIC ACID	ACETYLSALICYLIC ACID	ACETYLSALICYLIC ACID	ACETYLSALICYLIC ACID
Reported medication	ASPRIN			

In a table above, ASPRIN is coded to multiple ATC classes based on its action and properties. ASPRIN is used as Analgesic (pain killer, especially for headache), as antipyretic (reduce fever), as anti-inflammatory (reduce swelling) and also as anti-platelet (prevent blood clot).

More details about the WHODrug dictionary can be found at links mentioned in recommended reading section.

STANDARDIZED MEDDRA QUERY (SMQ)

In addition to the classification of events observed through MedDRA dictionary for the statistical analysis, there also exists a Standardized MedDRA Query (SMQ). SMQs groups the preferred terms (PTs) according to the medical condition the event falls under.

Table 3.1: PTs belong to different SOCs but the same SMQ.

SMQ Level 1 Name	Hypertension (SMQ)	Hypertension (SMQ)	Hypertension (SMQ)
Body System Organ Class (SOC)	Cardiac disorders	Metabolism and nutrition disorders	Investigations
Preferred Term (PT)	Hypertensive heart disease	Metabolic syndrome	Blood pressure increased
Reported AE Term	Hypertensive heart disease without heart failure, benign	Metabolic syndrome	blood pressure elevation

For the sake of example, other levels from MedDRA (LLT, HLT, HLGT) are not shown in the above table.

The reported Adverse events (AEs)/ Preferred terms (PTs) are classified to different SOCs per MedDRA dictionary structure. However, all of these PTs fall under same medical condition called Hypertension.

Points to note:

- Definition: SMQs are groupings of MedDRA PTs from multiple SOCs relating to a defined medical condition or area of interest.
- They are intended to aid identification of potentially relevant reports/ cases.
- MedDRA dictionary versions released twice in a year (1st of March and 1st of September). SMQs are a part of MedDRA releases and are updated concurrently to the MedDRA terminology.
- SMQs are organized in general (Level 1) to specific (Level 5) terms. Preferred terms may not have all SMQ levels but should be linked to at least general broad level of SMQ (i.e. Level 1). [As shown in the table below – PTs in last two columns do not have all SMQ levels]
- Not necessarily all preferred terms in MedDRA have its SMQ levels.
- Each of the SMQ terms is associated with the SMQ codes and the term codes. Term codes in SMQ are linked with lowest level term codes (LLT code) in adverse events dataset to retrieve SMQs for MedDRA Preferred terms (PTs).
- Two reference datasets SMQ LIST and SMQ CONTENT. LIST dataset contains SMQ name, code and corresponding level while CONTENT holds SMQ code, term code and additional info like status (active or inactive), scope (Broad and narrow search) and algorithm.

Table 3.2: Example of PTs with SMQ levels

SMQ Level 1 Name	Hepatic disorders (SMQ)	Hepatic disorders (SMQ)	Malignancies (SMQ)	Hypertension (SMQ)
SMQ Level 2 Name	Drug related hepatic disorders - comprehensive search (SMQ)	Drug related hepatic disorders - comprehensive search (SMQ)	Malignant or unspecified tumours (SMQ)	
SMQ Level 3 Name	Drug related hepatic disorders - severe events only (SMQ)	Drug related hepatic disorders - severe events only (SMQ)	Malignant tumours (SMQ)	
SMQ Level 4 Name	Liver neoplasms, malignant and unspecified (SMQ)	Liver neoplasms, malignant and unspecified (SMQ)		
SMQ Level 5 Name	Liver tumors of unspecified malignancy (SMQ)	Liver malignant tumors (SMQ)		
Preferred Term (PT)	Hepatic neoplasm	Hepatic neoplasm malignant	THYROID CANCER	BLOOD PRESSURE INCREASED
Reported Term	SUSPICION OF HEPATIC NODULA	LIVER EXPLANT-MULTIPLE HCCS	Thyroid Nodule 1.2 cm-malignant	elevated blood pressure

Additional details about SMQ's detailed structure, the search methods, can be found at links mentioned in recommended reading section.

UNIT CONVERSIONS

In clinical trials, it is commonly observed that the data about patients, for example: laboratory, vital signs are collected/ measured in different units at different investigational sites, because of the different conventions followed at sites or laboratories. For instance: In the US, weight is measured in pounds (LB) and Rest of the world measures it in Kilograms (kgs) and the corresponding reference range values from individual sites would also be different. However, for the analysis to be meaningful, all results and ranges must be described in the same unit for each test. This transformation of converting values in different units to the preferred units (US, SI) does need conversion factors for each of these labs, ECGs, vital signs tests. These conversion factors are stored and maintained by the individual companies within their data libraries unlike the dictionaries which are maintained by a separate organization.

Table 4.1.1: Simple vital signs example to depict the unit conversion

Site	Patient	Vital sign test	Result value	Result Unit	Standard result (in SI)	Standard Unit (SI)
101	1011	HEIGHT	67	IN	170.2	cm
101	1011	WEIGHT	193	LB	87.5	kg
102	1021	HEIGHT	62	IN	157.5	cm
102	1021	WEIGHT	55	Kg	55	kg

In the above table, two subjects from different investigational sites, have their vital signs data (Height, Weight) collected in different units but are converted to preferred units i.e. Standard International (SI).

The conversion factors for each of the tests and the decimal places up to which the converted values to be shown are the two requisites for performing conversions of any measured values. This information is generally maintained in two reference datasets named CONVERSION and PRECISION for all the Lab parameters, Vitals and ECG tests by individual pharmaceutical companies.

Table 4.1.2: CONVERSION

Parameter	Original units type	Pre-conversion Unit	Conversion unit type	Post-Conversion Unit	Conversion factor processor	Conversion factor
WEIGHT	SI	kg	SI	Kg	MULT	1
WEIGHT	SI	kg	US	LB	MULT	2.204622
WEIGHT	US	LB	US	LB	MULT	1
WEIGHT	US	LB	SI	Kg	MULT	0.453592

Table 4.1.3: PRECISION

Parameter	Unit	Precision Type	Precision Number
WEIGHT	LB	DECIMAL PLACES	1
WEIGHT	kg	DECIMAL PLACES	1
HEIGHT	IN	DECIMAL PLACES	0
HEIGHT	cm	DECIMAL PLACES	1

As you see in the table 4.1, the converted values in SI units for Height (in cm) and Weight (in kg) are shown up to decimal places 1 only based on the precision number for the respective parameter and unit (shown in table 4.1.3).

Table 4.2: Lab data example with reference ranges

Lab test	Result in original units	Original unit	Result in standard units	Standard unit	Ref. range upper limit in orig. units	Ref. range lower limit in orig. units	Ref. range upper limit in std. units	Ref. range lower limit in std. units	Gender
HGB	15.1	g/dL	151	g/L	17.7	12.6	177	126	Male
HGB	148	g/L	148	g/L	177	126	177	126	Male
HGB	148	g/L	148	g/L	162	116	162	116	Female
HGB	163	g/L	163	g/L	162	116	162	116	Female
HGB	168	g/L	168	g/L	178	129	178	129	Male

Lab parameter Hemoglobin (HGB) with both results and reference ranges in different units, but converted to preferred units of Standard International (SI).

Points to note:

- Reference ranges helps to flag abnormal values.
  [Example: Hemoglobin 163 is out of 116 to 162 in table 4.2]
- Reference range of parameters varies for different Age groups, Gender, Geographical regions, local conventions at site etc.
  [Example: Hemoglobin ranges are different for Male and Female patients and also within male patients in table 4.2]
- Precision is critical, as a small change (rounding decimals) could lead to wrong interpretation in certain instances.

CLINICALLY NOTABLE ABNORMALITIES

In the above example (table 4.2), we have learnt about the reference ranges [the Upper Limit of Normal [ULN] and Lower Limit of Normal (LLN)], through which the resulting value of the lab test is identified to be normal i.e. within the range or abnormal i.e. out of range. Further to this, we would also be interested to see if this abnormality in lab values is clinically significant or not and that could potentially be an adverse event or risk. So, there needs to be a set of criteria defined for each of the laboratory tests to identify the notable values called clinically notable abnormality criterions (as shown in below tables with Hemoglobin as example).

Table 4.3.1: Clinically notable abnormalities for lab parameters

Lab test	Mild (Grade 1)	Moderate (Grade 2)	Severe (Grade 3)	Potentially Life Threatening (Grade 4)
Hemoglobin in g/dL (Female)	11.0 - 12.0	9.5 - 10.9	8.0 - 9.4	<8.0
Hemoglobin in g/dL (Male)	12.5 - 13.5	10.5 - 12.4	8.5 - 10.4	<8.5

The notable abnormality ranges for male and female for LAB parameter hemoglobin in g/dL.

These criteria are not standard for all the trials. They are specific to study and defined in the protocol considering the disease/ therapeutic area, participants in a trial etc. and to what granularity a parameter to be analyzed.

For example: Table 4.3.2: Clinically notable abnormalities for lab parameters

Lab test	Notable Abnormality value
Hemoglobin in g/dL (< 16 years of age)	>=2 g/dL decrease from baseline (or) <8.5 g/dL
Hemoglobin in g/dL (>= 16 years of age)	>=2 g/dL decrease from baseline (or) <10 g/dL

The notable abnormality is defined for age groups and only one criterion in reference to baseline unlike the example in table 4.3.1 where there are several ranges defined for gender.

Let's see a simple example on vital sign tests to understand why the grade ranges for a parameter could be different from trial to trial.

Table 4.3.3: Normal and clinically notable abnormal values for vital sign tests

Vital sign test	Normal (mm/Hg)	Notable abnormal (Lower) (mm/Hg)	Notable abnormal (Upper) (mm/Hg)
Systolic Blood Pressure (SBP)	<120	90	140
Diastolic Blood Pressure (DBP)	<80	60	90

Above table 4.3.3, consists of normal and clinically notable abnormal values of SBP (Systolic Blood Pressure) and DBP (Diastolic Blood Pressure) in healthy or non-hypertensive person. However, if the trial is for hypertensive patients, the notable abnormality ranges or values in the above table are not justifiable. So, the ranges for such trial could completely be different to conduct analysis (example shown in table 4.3.4).

Table 4.3.4: Normal and clinically notable abnormal values for vital sign tests (for Hypertensive patients)

Vital sign test	Normal (mm/Hg)	Notable abnormal (Lower) (mm/Hg)	Notable abnormal (Upper) (mm/Hg)
Systolic Blood Pressure (SBP)	>=120	either <=90 + decrease >=30 or < 75	either >=180 + increase >=30 or > 200
Diastolic Blood Pressure (DBP)	>=80	either <=50 + decrease >=20 or < 40	either >=105 + increase >=20 or > 115

The above abnormal ranges are relative to baseline.

Points to note:

- Notable abnormalities should be defined in the protocol by clinical and safety experts.
- Not all lab parameters, vital signs or ECG tests are analyzed for notable abnormality.
- The notable abnormality criterions defined in the study protocol could base from the general toxicity criterions defined in the CTCAE (Common Terminology Criteria for Adverse Events) by NCI (National Cancer Institute).
  [For example: Below notable ranges for Hemoglobin have been referenced from CTCAE table located at http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf]

Table 4.3.5: Clinically notable abnormalities for lab parameters from CTCAE

Lab test	Mild (Grade 1)	Moderate (Grade 2)	Severe (Grade 3)	Potentially Life Threatening (Grade 4)	Fatal Adverse Event Grade 5
Hemoglobin (in g/dL)	<LLN – 10.0	<10.0 – 8.0	<8.0 – 6.5	<6.5	Death

The table above only shows ranges in one unit for the sake of simplicity.

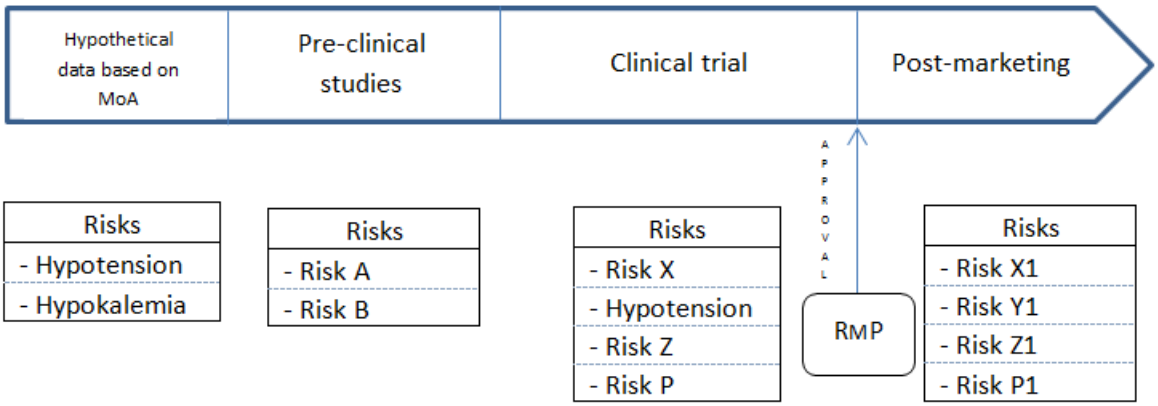
SAFETY CONCERNS SPECIFIC TO DRUG

With respect to the safety concerns of a drug, it is important to monitor adverse events which can also include elevated/ abnormal laboratory values as noted in the above section. Based on the information from past conducted trials, published literature reports, epidemiology data, preclinical data there is already an idea about the safety profile of the drug and some of the observed AEs or SAEs in an ongoing clinical trial that could potentially raise a safety concerns and needs to be reported distinctly, are called AEs of special interest. In addition to special AEs, we also need to identify and report the risks associated with the drug, and assess the benefit-risk profile.

For example: Subject experiencing an adverse effect of alopecia who is receiving a drug for acne. This is potentially a risk that needs to be completely evaluated. However, if a patient develops alopecia while receiving an anti-neoplastic drug, the importance of alopecia will not be as high as in the patient who received it for acne.

Identification of risks is an extensive and continuous exercise that could base from the literature, the mechanism of action, an epidemiology data, publications, competitor's data, from pre-clinical studies through the clinical trial as well as post marketing. The risks identified all through and that are anticipated will need to be reported, this called Risk Management Plan (RMP).

For example: Suppose, we are conducting a trial on assessing the safety and efficacy of a new anti-hypertensive drug.



Points to note:

- Risks are generally categorized into – Identified risks, Potential risks and Missing information.
  - In the example above, Hypotension and Hypokalemia are expected to be risks based on the mechanism of action (MoA) of Diuretics. Risks A and B have been identified in pre-clinical studies. Likewise, additional risks are identified from the epidemiology studies, literature, publications, and our ongoing clinical trial. All these risks could be characterized as either potential or identified. [Example: If a risk has been identified in only 10 out of 1000 patients, it may not be considered as identified risk but considered to be a potential risk]
  - Hypokalemia was anticipated to be a risk based on MoA but it was not observed in our clinical trial – this may be considered as a potential risk. This is because theoretically, the patients could develop hypokalemia because of the way the drug acts, and more data from other sources may be needed for further evaluation.
  - If a drug is metabolized in the liver, then the drug may act differently in a patient with liver failure. If this information is not obtained from clinical trials (difference in drug action in patients with liver failure), then it becomes a topic of missing information.
- Risk evaluation is a continuous process of a drug life cycle.
- Not all AEs/ SAEs would be risk (Alopecia example discussed above).
- RMP is to specify what is known (identified risks and potential risks) and what is not known (missing information) about the safety of a drug at the time of submission for approval.
- AEs of special interest and Risks are provided by Drug safety team.

EXTERNAL/ OPEN SOURCE DATA

This section of the paper illustrates the usage of external information for the clinical trial data analysis, which could come from the previously conducted trials or general surveys i.e. from publications or literature etc. available on the internet.

Let's consider an example; we have to provide an answer to below health authority question.

"Whether the death rate in our trial is more, or less than what is observed in the literature?"

A few details about the trial:

- We have conducted a clinical trial using treatments active drug "Ta" and Placebo "Tp" on patients with disease "Da".
- It is tested on patients with age above say 55 as the disease "Da" is mostly observed in elders.
- The planned analysis for death rate is by age groups say 55-<65, 65-<75, 75-<85 and >=85 and by gender.

To answer the above question, we need below information.

i. First, we need the death rate observed in general population.

Table 4.1: Death rate observed in general population.

Category	Death rate observed in general population (Males) (Per 1000)	Death rate observed in general population (Females) (Per 1000)
55-<65	10.9	6.5
65-<75	23.5	15.8
75-<85	58.7	42.3
>=85	155.7	123.1

Note: The numbers shown in the table are hypothetical for illustration.

Note: The above death rates in general population as per planned analysis (i.e. by age groups and by gender) are derived from the CDC (Centers for Disease Control) Life table US, available on the internet. Statistician would get these death rates calculated for programmers. In the CDC Life table, the death occurrences of a hypothetical cohort of 100,000 persons, per age are given (snippet shown table 4.2 for reference).

Table 4.2: CDC Life Table US

Table 1. Life table for the total population: United States, 2010

Spreadsheet version available from: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/63_07/T

Age	Probability of dying between ages x to x+1 q(x)	Number surviving to age x l(x)	Number dying between ages x to x+1 d(x)
0-1	0.006123	100,000	612
1-2	0.000428	99,388	43
2-3	0.000275	99,345	27
3-4	0.000211	99,318	21
4-5	0.000158	99,297	16
5-6	0.000145	99,281	14
…..			
…..			
92-93	0.172661	17,061	2,946
93-94	0.188988	14,115	2,668
94-95	0.206214	11,447	2,361
95-96	0.224274	9,087	2,038
96-97	0.243080	7,049	1,713
97-98	0.262527	5,335	1,401
98-99	0.282492	3,935	1,112
99-100	0.302838	2,823	855
100 and over	1.000000	1,968	1,968

Note: This table gets updated every year or so. Pick the data from recent version of the table.

ii. Second, we need to calculate the death rate in our trial for both active “Ta” and Placebo “Tp” treatment groups.

Table 4.3: Death rates observed in our trial for both active and placebo treatments.

Category	Death rate observed in Active treatment [Ta] (Males) (Per 1000)	Death rate observed in Active treatment [Ta] (Females) (Per 1000)	Death rate observed in Placebo treatment [Tp] (Males) (Per 1000)	Death rate observed in Placebo treatment [Tp] (Females) (Per 1000)
55-<65	8.6	5.8	9.8	6.1
65-<75	21.2	14.9	23.4	15.5
75-<85	58.8	42.1	59.4	43.1
>=85	155.4	125.5	158.4	126.1

Note: The numbers shown in the table are hypothetical for illustration.

Inferences:

1) If the death rate observed in the active group is less than that of what is observed in the literature (i.e. above Table 4.3), we can claim our drug is doing well with respect to death occurrences when compared to the death occurrences in general population.

2) In addition, if the death rate observed in the placebo group is close to what is observed in the literature (i.e. Table 4.3), we can claim our trial conduct was good, which further strengthens the first point. (No biased sample etc.)

However, suppose if there is already an article or publication available (from the past trial conducts or surveys) on the death rates among the patients with disease "Da", then one would use this data to compare, as this is more specific to the diseased patients of our interest than the death rates observed in the general population.

Note: The most relevant and recent data available in the literature or publications should be used for the analysis. These references/ sources will be provided by qualified Safety and Epidemiology personnel.

TRIAL DESIGN MODEL

In addition to the standard dictionaries and references, the evolution of standards in data collection and reporting (i.e. CDISC standards and its models) demands the trial design datasets to be generated with the trial specific details describing the aspects of the planned conduct of a trial for submission, which in turn could serve as a reference throughout the trial.

The datasets are a part of Trial design model (TDM) of CDISC and includes Trial Visits (TV), Trial Arms (TA), Trial Elements (TE), Trial Inclusion/Exclusion (TI) and Trial Summary (TS).

These standard trial design datasets provide a standardized way to describe the study trial and allows reviewers to:

- Clearly and quickly grasp the design of a clinical trial
- Compare the designs of different trials
- Search a data warehouse for clinical trials with certain features
- Compare planned and actual treatments and visits for subjects in a clinical trial.

Please refer to the SDTM Implementation guide for complete details.

CONCLUSION

In order to convert the CRF collected data into analytical form, we need to classify or standardize the data, identify patterns within data or compare against the past available or the reference data etc. using a wide variety of supplemental information. The Statistical analysis programmer should have an understanding of this supplementary data coming from different sources in a variety of formats. In this paper, we discussed typically used standard dictionaries, references and the information from safety, clinical, statistical and other teams. However, there could be several other supplementary data specific to the disease or therapeutic area.

REFERENCES

MedDRA: http://www.meddra.org/how-to-use/basics/hierarchy

CTCAE v3.0 table: http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf

CDC Life table US: ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/63_07/Table01.xlsx

ACKNOWLEDGMENTS

We would like to express our sincere thanks to fellow colleagues mentioned below for all the discussions, inputs and review on the content of this paper.
Alla Srinivas
Angshuman Sarkar
Ambroise Habersetzer
Boda MallaReddy
Kartik Rajan
Kiran Jayam
Samina Hussain

RECOMMENDED READING

Paper: Using the WHO Drug Dictionary for Reporting Clinical Trials
http://www.mwsug.org/proceedings/2007/stats/MWSUG-2007-S06.pdf

Introductory Guide for Standardised MedDRA Queries (SMQs) at:
http://www.meddra.org/sites/default/files/guidance/file/SMQ_intguide_15_1_English_0.pdf

Paper for SMQ search methods: “Everything You Need To Know About Standardised MedDRA Queries”

SDTM Implementation Guide
http://meta-x.com/cdisc/doc/SDTM%20Implementation%20Guide%20V3.1.2.pdf

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Author Name: Satyapal Gudla
Company: Novartis Healthcare Pvt. Ltd.
Address: Hi-tech city, Hyderabad, India
City / Postcode: Hyderabad
Email: Satyapal.gudla@novartis.com

Author Name: Nigel Montgomery
Company: Novartis Pharma AG
Address: Lichtstrasse, Basel, Switzerland
City / Postcode: Basel
Email: nigel.montgomery@novartis.com