General Output Tips and Considerations

phuse.eu

Contents

Doc ID: WP-034 | Version: 1.0 | Working Group: Standard Analyses and Code Sharing | Date: 03-Aug-2020

I. Overview: Purpose of this Document

This white paper provides general advice on the display of data in tabular form. The intent is educational and not prescriptive, as an aid to considering how to present data in tables. Tables allow for the organisation of alphanumeric data for easy reference and comparison.

Edward Tufte, an expert in the visual display of data, notes the limitation of presenting data in sentences:

> The conventional sentence is a poor way to show more than two numbers because it prevents comparisons within the data. The linearly organized flow of words, folded over at arbitrary points (decided not by content but by the happenstance of column width), offers less than one effective dimension for organizing the data. [1]

This guide is intended as a reference document to accompany other white papers describing different types of analyses. The displays in each paper may differ in style and presentation and are within the norms of standard outputs. They are not intended as firm requirements for the outputs as each company will add their own flavour and conventions to their presentations.

A. Scope

The scope of this guidance covers the general elements in the presentation of tables and listings for clinical study reports (CSRs) and other similar regulatory reports. Here we discuss what are the minimum necessary elements for each to provide sufficient documentation for source, purpose and display. The parts of a table are defined for ease of discussion. The scope does not go into actual programming code, nor does it describe a finite look and feel for tables or listings. Rather, general principles are provided with as much flexibility as possible for actual use by members of pharmaceutical companies, CROs or other service providers as they create deliverables for regulatory agencies.

While graphical displays are becoming more commonly used as primary displays, formatting considerations of the displays are out of the scope of this paper. Some useful references for graphical displays follow:

Duke, Susan P.; Bancken, Fabrice; Crowe, Brenda; Soukup, Mat; . Botsis, Taxiarchis and Forshee, Richard. Seeing is believing: good graphic design principles for medical research. *Statistics in Medicine*, 2015, no. 34 (22):3040-3059.

Lang, T.A. and Secic, M. *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors, and Reviewers*, 2nd edition. Philadelphia: American College of Physicians, 2006.

Statistics How To (web page). Types of graphs used in Math and Statistics. Available at: *https://www.statisticshowto.com/types-graphs/*. Accessed 3 April, 2020.

II. General Guidelines on the Tabular Display of Data

A. Parts of a Table

Below are definitions of table parts, followed by a graphic visualisation of where these parts belong on the output. The definitions are standard descriptive parts of a table taken from the reference sources listed.

Table 1: Parts of a Table

Part Name	Description
Title	The title is a caption of a table beginning with "Table" followed by sequenced numbers or letters and separated by full stops. The title includes additional lines describing the content.
Headings and Stubs	Column headings define values in the columns. Stubs define the values in the rows. A stub does not require a column heading. Stubs are to rows what headings are to columns.
Straddles	Straddles are a type of column heading that spans two or more columns.
Columns	The vertically aligned display of words, numbers or symbols.
Rows	The horizontally aligned display of words, numbers or symbols.
Body	The rows and columns displaying units of measure defined by column headings and usually row stubs. Stub headings along with column headings define the values in the rows.
Legends	Lines of text placed immediately after the body of the table containing statements explaining the use of terms or values. May be prefaced with a word such as 'Note:'.
Abbreviations and Definitions	Lines of text defining abbreviations, acronyms, and terms used in the body.
Footnotes	Placed after the abbreviations and definitions, if any. Footnotes lead with symbols (asterisk, dagger, double dagger, or section symbol) or superscripted letters ([a, b, c]) or numbers ([1, 2, 3]) which are in the table behind units of measure or words. Footnotes add information regarding the associated values where they occur in the body.

Source: APS Manual of Style – parts of a table. [2]

Figure 1 is a graphic representation of parts for a standard table. Note that the organisation of tables, particularly the order or location of legends, abbreviations and definitions, and footnotes may vary across pharmaceutical company standards. The main point is that the various elements described as needed are presented in your outputs.

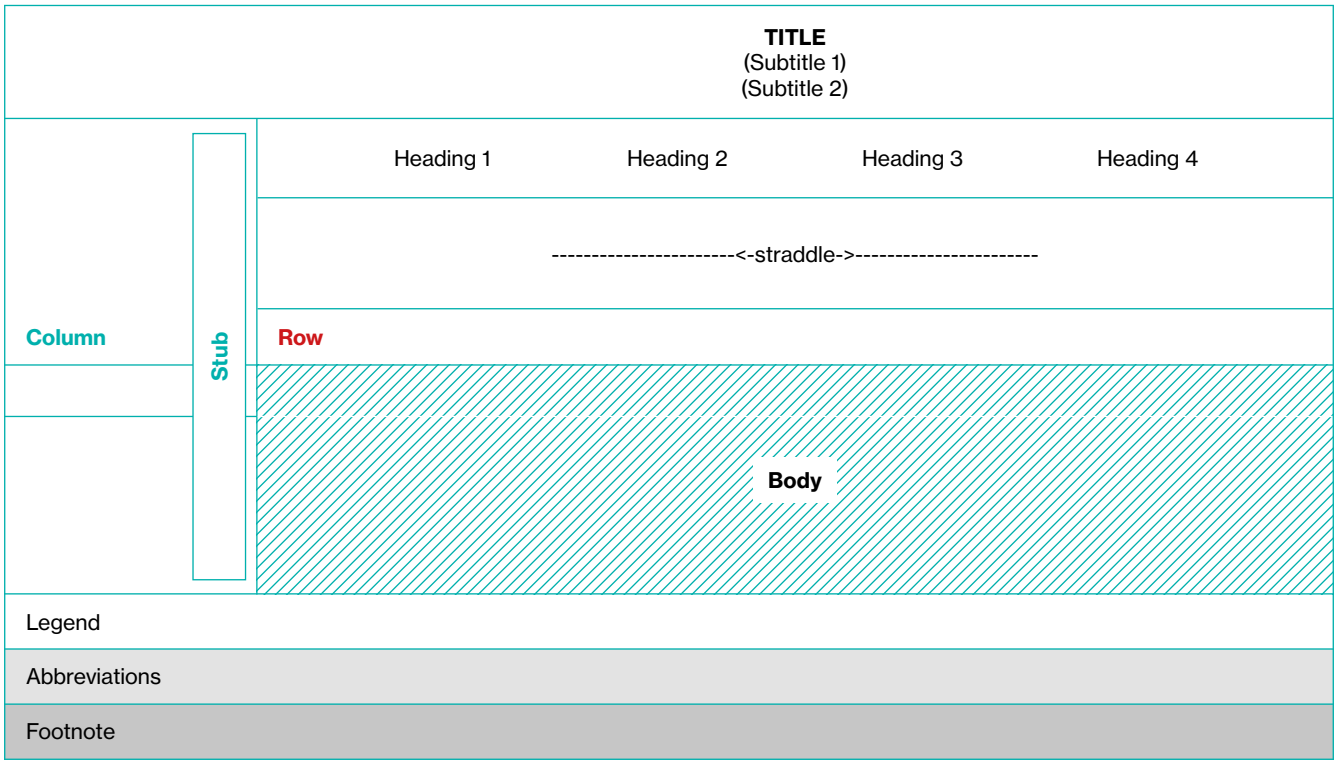


Figure 1: Parts of a Table

Now that we have defined the parts of a table they will be referenced by the terms as shown. We recommend following the order of the APS manual; however, alternative ordering of legends, abbreviations and footnotes is acceptable.

1. Title

The title of a table needs to clearly identify the contents. It begins on the first line with a table reference number. We recommend, although it is not a requirement, that the reference numbers follow the ICH-9 numbering conventions.

> ☞ At the minimum, titles should follow internal company standards so that they can be placed in the CSR or other document without having to be updated during publishing activities.

In the context of clinical trial data, we recommend that at a minimum, the first title line contain the description of the summary/analysis and the next line contain the population. In some cases, more than two lines will be needed for the description summary. See also Section VI, Note on Pharmaceutical Company Standards.

The table number and title must be a unique combination. This will aid in identifying a table in the report, as well as aid in generating titles for tables programmatically. For sub-tables, add .1 to the numbering sequence and a "by xxx" to the title to identify the sub-group being shown. When describing "by Treatment Group" some minimalists feel that if all the tables are by the same groups, this information can be omitted from the title. However, if you have multiple types of tables displaying various subgroups, such as "by Cohort", then this information needs to remain. (See examples in Section VI, Formatting and Rendering Issues.) Another phrase that often appears on table titles is "Summary of". This is usually the case for tables, so may also be left off.

> ☞ Note that the numbering sequence is generally not padded with 00s, which does offer a challenge in sorting. For example, table 14.1.12 comes after table 14.1.3. In your programming, you can account for this by adding an ordering variable as the first column, or using row numbers, making sure the table names are in the required order.

Table Title Examples:

The first example includes the study ID and description and may be either programmed in the table by drawing from a spreadsheet or added separately as a Word caption. Note that concatenating the table number, title and population saves space, allows the easy creation of a Word TOC for tables, and can be linked as a cross-reference in the text of the document.

Table ABCD.14.1. Demographic and Baseline Characteristics (ITT Population)

The second example is a traditional three-line minimum title which may be programmed as part of the table by drawing from a spreadsheet with multiple columns. Both examples are acceptable, although we recommend the former as a means to save space on the table for more data displayed in the body of the table. Use of abbreviations for population is allowed.

Table 14.1
Demographic and Baseline Characteristics
Intent-to-Treat Population

2. Headings and Stubs

The headings of a table generally list the treatment arms and comparators. These treatment groups should be clearly defined. If more than one dosing strategy is displayed, this information should be included rather than a generic "<treatment> 1", "<treatment> 2". As a convention, we recommend starting with Placebo on the left, then increasing dosage to the right, and finally an "All Treatment" column if multiple study drug treatments are shown. In addition, a column may be added for comparisons such as "Change from Baseline".

Stubs describe the contents of the row. The data is then read from left to right, with the column headers identifying the subset being described, along with associated summary statistics. The stubs may be multi-level, including such items as timepoint (analysis visit) or further sub-setting the display. In that case, the high-level stub might only have counts instead of counts and percentages. The stub is indented to visually assist the reader in viewing the data. (See Figure 1.)

Example of a Stub	
Overall	
Any Major Deviation	xx (xx.x%)
<Deviation #1>	xx (xx.x%)
<Deviation #2>	xx (xx.x%)

Sometimes there are too many treatment groups for a clear display. In that case, the data can be transposed. A transposed display shows the treatment groups as the stubs and the columns show the variables being described.

3. Body of the Table

The body of the table is the content and should fully support the title of the table. We recommend that, if applicable, all values displayed should be associated with a **unit of measure**, which is usually stated in the **stub** or **column head**, and sometimes in the table **legend**. If all units of measure in the table body are the same, they can be moved into the title or stubs, as appropriate.

4. Abbreviations, Legends and Footnotes

To the casual reader, these three items are one and the same, and they appear at the bottom of a table. Generally, the abbreviations come first to define acronyms or terms that might need explanation. The second element is a legend item, which is a group of sentences that applies to the overall table. This might be a description of statistical methods used, whether and how missing data was imputed, and types of p-values used. The third item is the footnote. This is referenced in the table usually with a superscript or subscript, and then defined in detail in the footnote. The order is flexible and is based on your company's process or preference.

One more line that appears below the footnote but is not really part of the footnote is information for traceability of the source program. Generally, this includes the name (and path) of the program, name of the dataset and date/time stamp of run. Traceability to the origins of content is a key component of a display.

☞ If the footnote takes up more than a third of the page, it is a good practice to wait until the last page of that output and put everything there. In this way you can reduce the size of the output by increasing the available space for the body. Additionally, abbreviations or definitions may be added to each page, but footnotes left to the end. A third method is to only reference the footnotes that appear on that page, but that becomes an operational challenge for programming.

III. Types of Tables

A. General Types of Tables

The types of tables listed below are used in analyses depending on structure of the data and the needs for the research. These types are not explicitly part of the table title, but it is good practice to note so that we can describe further the purposes of each.

The layout of a table determines how elements of the table are defined. An "item" here is the thing being measured or referenced. An "attribute" is the measure or value for the item.

- If the items are named in the table stub and attributes are named in the table headings, the layout is "horizontal".
- If the items are named in the table headings and the attributes are named in the table stub, the layout is "vertical".

Table 2: Definitions for Types of Tables

Category	Items are listed in the stub with subsequent columns showing attributes of the items, as named in the column headings Example: Column 1 lists subject/patient numbers Columns 2–5 list doses received at each visit
Comparison	Attributes are listed in the stub with subsequent columns showing values for different items Example: Column 1 lists age, weight, height, gender, etc. Columns 1–5 list corresponding values for each subject/patient, as named in the column headings
Analysis	A table containing data and the results of statistical analysis of data, or summary statistics.
Listing	A table presenting data without analyses.

B. Medical and Pharmaceutical Research Tables for Inclusion in a Clinical Study Report (CSR)

Statistical tables show the values of the cumulative distribution, probability, or probability density functions of certain common distributions for different values of their parameters.

Table 3: Definitions for Content

Demographics and Baseline Characteristics	Study population dispositions such as age, weight, gender, laboratory measures, disease state, concomitant medications, treatment compliance, discontinuations. Usually reported across study visits.
Safety	Study results for treatment emergent adverse events, serious adverse events, deaths, overdose, particular safety values for particular body systems such as for cardiac, glucose, liver enzymes, laboratory values, vital signs, immunogenicity.
Efficacy	Analyses of endpoint results, changes from baseline, per visit and per patient changes, drug dosage and responses, pharmacokinetic/ pharmacodynamic analyses, etc.
Tabulation	Complete listing of collected data by subject or another identifier.

IV. Considerations for Organising Data in a Table

A. Note on Pharmaceutical Company Standards

In general, targeted data, analyses, and the format of statistical output intended for presentation in regulatory documents are defined in collaboration between the project statistician, the statistical analyst, the medical communications writer, the clinical research physician or scientist, and other members of the writing team. Planning of analyses and formatting may include the collaboration of partners and/or third-party operators.

> ☞ A good rule to follow is that tabular material must stand alone, therefore all symbols and abbreviations used should be defined at first use. [2]

Company standards for statistical output will vary. It is important to be consistent. For example:

- Table numbering may be pre-planned and programmatically added in the table or,
- table numbering may be added above the table as styled Word text or Word table captions for auto-numbering and cross-referencing.
- The description of the table may be after the table number or begin on the second line.
- Some companies may number tables in a single sequence from the beginning of the document, e.g. Table **1**, Table **2** or,
- the numbering sequence may include Level 1 and lower section numbers followed by a sequence number for the individual table, e.g. with Level 1, Table **14**.1, Table **14**.2, etc.; and with Level 1 and Level 2, Table **14.3**.1, Table **14.3**.2, etc.
- Appendix and Attachment sections may use table numbering that denotes their position, e.g. Table APP.10.1.

> ☞ A company should create and maintain standards for statistical output.

V. Regulatory Requirements

A. ICH Guidelines

The International Council for Harmonisation (ICH) provides guidance which was generally adopted by the regulatory agencies of the FDA, PMDA, EMA and Health Canada. It is important to note that the current document should be used as a guidance and the recommendations are not mandated by any regulatory agency. They are provided to assist the sponsor company with making decisions regarding their submission process. The ICH E3 Guideline for Industry: Structure and Content of Clinical Trial Study Reports was last updated in 1996 and still holds today. [3, 4] This guideline provides some basic assumptions that are recommended for outputs. The guideline is referred to by the FDA on their website, ICH Guidance Documents. [5, 6]

The documents presented there are for safety, efficacy, quality and joint safety/efficacy (multidisciplinary). The ICH E3 talks mostly about content that should be presented in the body of the report, and in general terms, when tabular output should be used. For example, on page 32 it describes how to present a tabular listing of laboratory findings. Additional guidance is given to break up the data into logical parts so that if the data does not fit onto one page, such as if not all tests can be displayed in a single table, they should be grouped logically.

In addition, a new document was issued in 2013 to clarify some of the content in the ICH E3. This is the Guidance for Industry E3 Structure and Content of Clinical Study Reports Questions and Answers (R1). [4]

ICH Guidance also provides table numbering guidelines. Although these are not required, they are recommended as a starting point for numbering your CSR. The basic guideline states where the various sections, such as demographics, safety and efficacy, should be organised in the CSR.

The EMA also supports the use of ICH E3 Guidance. We need to be aware of the privacy requirements for each country as they vary and note that the EMA will publish study reports. Do check for any additional requirements for other countries not in the EMEA such as Japan, Eastern Europe and South America.

> ☞ We highly recommend avoiding any obvious errors in grammar, punctuation and spelling by running appropriate checks on the final document.

VI. Formatting and Rendering Issues

1. Column Formatting in Microsoft Word

Columns in tables in Word-compatible formats may be aligned using a monospaced font or defined by columns in a Word table. A table in Word created using a monospaced font for spacing is, in this document, called a text table. An example of a monospaced font table using Courier New font:

```
Type      Colour    Gender
Cat       Tabby     Male
Dog       Brown     Male
```

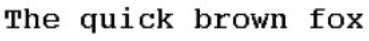
A Word table is a Word object with columns, rows, borders, and other properties such as shading, border style, colour, repeating header and table style. Auto-formatting may be enabled in Word tables including auto-fit to contents, visible window, or allow breaking across pages.

A Word table has more attributes to control and, when inserted into a Word document in a co-authoring environment, will impose a greater burden on the system than text tables. All tables with borders in this document are Word tables.

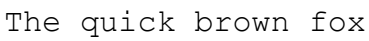
2. Table Fonts

Table fonts should be scalable, that is, on an electronic display the fonts will not pixelate when the view is zoomed in. Scalable fonts are called vector or outline fonts and are most commonly True Type or Microsoft OpenType fonts. A non-scalable font is called a bitmap or raster font. Examples:

Courier Regular is a bitmap font and will pixelate on zooming in:



Courier New is a TrueType vector font and will not pixelate on zooming in.



3. File Formats for Tables

Output table file formats may be:

- plain text .lst or .txt
- Word-related .odt, .rtf, .doc. or .docx
- spreadsheets viewable in Excel such .csv, .xls, .xlsx, .xml, .html
- formats viewable via Acrobat tools such as .pdf or .eps, or
- proprietary formats unique to the production tools.

Refer to your individual company's practices for using output file formats. Keep in mind file size can be an issue in some systems.

How statistical output is integrated into a regulatory submission document will also vary by company. Some output may be inserted into documents during authoring in Word or by another program. Other systems may keep output separate from the document content until the document is assembled at finalisation, usually in PDF format. For some companies, statistical output may be predominantly relegated to appendices and not incorporated into the body of a regulatory document.

VII. Programming Tips

Companies may vary in their table numberings, where they may use sub-headings or sub-levels. This may be dictated by internal programming processes, and how the table titles are applied. Table numbering may be programmed directly or may be read in from an external file.

> ☞ To automate the process and simplify the QC, we recommend creating numbering, titles and footnotes outside of the program so that they may be updated in the metadata rather than in each program. This may serve a dual purpose of tracking the progress.

VIII. Disclaimer

The opinions expressed in this document are those of the authors and should not be construed to represent the opinions of PHUSE members' respective companies or organisations or the FDA's views or policies. The content in this document should not be interpreted as a data standard and/or information required by regulatory authorities.

IX. References and Resources

[1] Tufte, Edward, *The Visual Display of Quantitative Information*, Cheshire, Conn. Graphics Press, 2001, p.178.

[2] Parts of a Table, APS Style Manual. Available at: *web.archive.org/web/20170116080958/http://www.apsstylemanual.org/oldmanual/parts/text/tables.htm*. Accessed 18 April, 2019 (Archived April 2020).

[3] [EMA] European Medicines Agency. Note for Guidance on Structure and Content of Clinical Study Reports. CPMP/ICH/137/1995. Step 5. Available at: *https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e-3-structure-content-clinical-study-reports-step-5_en.pdf*. Accessed 30 March, 2020.

[4] [EMA] European Medicines Agency. ICH E3 Questions and Answers. January 2012. Available at *https://www.fda.gov/media/84857/download*. Accessed 30 March, 2020.

[5] [FDA] United States Food and Drug Administration. E3 Structure and Content of Clinical Study Reports. July 1996. Available at: *https://www.fda.gov/regulatory-information/search-fda-guidance-documents/e3-structure-and-content-clinical-study-reports*. Accessed 30 March, 2020.

[6] ICH. Quality, Safety, Efficacy, and Multidisciplinary Guidelines. Available at: *https://www.ich.org/page/ich-guidelines*. Accessed 17 March, 2020.

X. Project Contact Information

Name: Mary Nilsson
(Analysis and Display White Papers Project Team Leader)
Enterprise: Eli Lilly & Company
City, State ZIP: Indianapolis, IN 46285
Work Phone: 317-651-8041
E-mail: *nilsson_mary_e@lilly.com*

XI. Acknowledgements

The primary contributors include Karin A. LaPann and Mark E. Rohe. We thank additional contributors/reviewers (who participated in discussions and provided review comments on portions of the content) including Chris Battison, Nancy Brucken, Jim Cawley, Ye Fang, Sunil Gupta, Matthew Jones, Jane Liao, Kim Musgrave, Mary Nilsson, Mitra Pranab Kumar, Alan Shapiro, Lori Van Meter, Olga Vovk, Jerry Xu and Aiming Yang. Apologies to anyone we may have missed.