INTERNATIONAL COUNCIL FOR HARMONISATION OF
TECHNICAL REQUIREMENTS FOR PHARMACEUTICALS FOR HUMAN USE

ICH M8 Expert Working Group

Specification for Submission Formats for eCTD

v1.1

November 10, 2016

DOCUMENT CHANGE HISTORY

Version	Date	Comments
1.0	10 December 2015	Initial Step 4 document.
1.1	10 November 2016	Revisions based on M8 Review and the following change requests: 00020, 00030 and 00050.

Table of Contents

1. INTRODUCTION

This additional specification describes the way files should be constructed for inclusion in the eCTD. This section includes file formats that are commonly used in electronic submissions. Other formats can be used according to guidance published in each region.

> The content in this document should be used in conjunction with the Regional/Module 1 Implementation Guides for additional information.

1.1 PDF

Adobe Portable Document Format (PDF) is a published format compliant to ISO 32000-1:2008. It is not necessary to use a product from Adobe or from any specific company to produce PDF documents. PDF is accepted as a standard for documents defined in this specification. The following recommendations support the creation of PDF files that regulatory authorities can review effectively.

To ensure that PDF files can be accessed efficiently, optimize PDF files for fast web view.

1.1.1 Version

All ICH regional regulatory authorities are able to read and accept PDF files saved as PDF version 1.4 through 1.7, PDF/A-1, or PDF/A-2 compliant to ISO 32000-1:2008. Regulatory authorities should not need any additional software to read and navigate the PDF files.

1.1.2 File Size

The size of a PDF file should not exceed 500MB.

1.1.3 Fonts

PDF viewing software automatically substitutes a font to display text if the font used to create the text is unavailable on the reviewer's computer. Font substitution can affect a document's appearance and structure, and, in some cases, the view of the document contents may be affected. Regulatory authorities cannot guarantee the availability of any fonts except Times New Roman, Arial, and Courier and fonts supported in the Acrobat product set itself. Therefore, all additional fonts used in the PDF files should be embedded to ensure that those fonts would always be available to the reviewer. When embedding fonts, it is suitable to embed a subset of the font. A subset includes only those characters used in the document and does not need to include all characters in the font type. All two-byte fonts, such as Japanese characters, should be embedded as a subset. When using Japanese characters, rules of operation should be established between the sender and receiver.

Although the embedding of fonts requires additional computer storage space, there are three techniques to help limit the storage space taken by embedding fonts include:

- Limiting the number of fonts used in each document

- Using only True Type or Adobe Type 1 fonts

- Avoiding customized fonts.

The following should be considered when embedding fonts:

Advantages:

- Embedding fonts allows the PDF file to be correctly displayed and printed on any receiving PC environment.

- The computer does not need the original fonts installed.

Disadvantages:

- The file size increases when fonts are embedded.

- When document contains many pages, this can make the document slower to print.

- Many eCTD documents contain a large number of pages. Printing time in such cases becomes a concern.

1.1.4 Font Size

Resizing a document because the contents are too small to read is inefficient. Times New Roman, 12-point font, the font used for this document, is adequate in size for narrative text and should be used whenever possible. It is sometimes tempting to use fonts which are smaller than 12 point in tables and charts but this should be avoided whenever possible.

When choosing a font size for tables, a balance should be sought between providing sufficient information on a single page to facilitate data comparisons for the reviewer while maintaining a font size that remains legible. Generally, Times New Roman font sizes 9-10 or an equivalent size of other recommended fonts are considered acceptable in tables but smaller font sizes should be avoided. The corollary of this is that in using larger font size, more tables might be necessary, which can complicate data comparisons since data might now be included in separate tables.

1.1.5 Use of Color Fonts

The use of a black font color is recommended. Blue can be used for hypertext links. Light colors can be difficult to read on a monitor as well as when printed, and should be avoided. The use of background shading can be difficult to read and should be avoided.

1.1.6 Page Orientation

Pages should be properly oriented so that all portrait pages are presented in portrait and all landscape pages are presented in landscape. To achieve this, the page orientation of landscape pages should be set to landscape prior to saving the PDF document in final form.

1.1.7 Page Size and Margins

The print area for pages should fit on a sheet of A4 (210 x 297 mm) and Letter (8.5" x 11") paper. A sufficient margin of at least 2.0 cm on the left side of each page for portrait and top of the page for landscape should be provided to avoid obscuring information if the reviewer subsequently prints and binds the pages for temporary use. The remaining margins should be a minimum of 0.8 cm. Header and footer information can appear within these margins but should not appear so close to the page edge to risk being lost upon printing.

1.1.8 Headers and Footers

The M4 Granularity document specifies that all pages of a document should include a unique header or footer that briefly identifies its subject matter. With the eCTD there is a significant amount of metadata available to the reviewer to allow easy identification of the document but it is still appropriate to have a unique identifier on each page (header or footer) of the document (e.g., when the document is printed or multiple documents are viewed on screen at the same time). The unique identifier does not necessarily have to contain the CTD section identifier or other metadata. It should be sufficient to identify the general subject matter of the document (e.g., study identifier, batch number).

1.1.9 Source of Electronic Document

PDF documents produced by scanning paper documents are usually inferior to those produced from an electronic source document. Scanned documents saved as image files are more difficult to read and do not allow reviewers to search or copy and paste text for editing. Scanning should be avoided where possible.

1.1.10 Methods for Creating PDF Documents and Images

The method used for creating PDF documents should produce the best replication of a paper document. To ensure that the paper and PDF version of the document are the same, the document should be printed from the PDF version. Documents that are available only in paper should be scanned at resolutions that will ensure the pages are legible both on the computer screen and when printed. At the same time, the file size should be limited. It is recommended that scanning be undertaken at a resolution of 300 dots per inch (dpi) to balance legibility and file size. The use of grayscale or color is discouraged because of file size. After scanning, resampling to a lower resolution should be avoided.

When creating PDF files containing images, the images should not be downsampled. Downsampling does not preserve all of the pixels in the original. For PDF images, one of the following lossless compression techniques should be used:

- For lossless compression of color and grayscale images, use Zip/Flate (one technique with two names). This is specified in Internet RFC 1950 and RFC 1951.

- For lossless compression of black and white images, use the CCITT Group 4 Fax compression technique. It is specified as CCITT recommendations T.6 (1988) - *Facsimile coding schemes and coding control functions for Group 4 facsimile apparatus.*

Paper documents containing hand-written notes should be scanned at a resolution of at least 300 dpi.

Hand-written notes should be done in black ink for clarity. Higher resolution is specifically requested when scanning documents containing non-Western characters (e.g., Kanji); 600 dpi is recommended.

For photographs, the image should be obtained with a resolution of 600 dpi. If black and white photos are submitted, 8-bit grayscale images should be considered. If color photos are submitted, 24-bit RGB images should be considered. A captured image should not be subjected to non-uniform scaling (i.e., sizing). Gels and karyotypes should be scanned directly, rather than from

photographs. Scanning should be at 600 dpi and 8-bit grayscale depth. Plotter output graphics should be scanned or captured digitally at 300 dpi.

High-pressure liquid chromatography or similar images should be scanned at a resolution of at least 300 dpi.

Applicants should validate the quality of the renditions.

1.1.11   Hypertext Linking and Bookmarks

Hypertext links and bookmarks improve navigation through PDF documents. Hypertext links can be designated by rectangles using thin lines or by blue text as appropriate. Bookmarks are expected even if there is no TOC In the document. A hypertext linked TOC and bookmarks should be included in documents 5 pages or longer. Literature References are the exception to these recommendations as the files may be protected and cannot be modified.

In general, for documents with a table of contents, bookmarks for each item listed in the table of contents should be provided. Bookmarks should include all tables, figures, publications, other references, and appendices even if these items are not in the table of contents. These bookmarks are essential for the efficient navigation through documents. The use of no more than 4 levels in the hierarchy is recommended, but additional levels could be created for study reports if such bookmarks contribute to efficient navigation.

Hypertext links throughout the document to support annotations, related sections, references, appendices, tables, or figures that are not located on the same page are helpful and improve navigation efficiency.

Relative paths should be used when creating hypertext links to minimize the loss of hyperlink functionality when folders are moved between disk drives. Absolute links that reference specific drives and root directories will no longer work once the submission is loaded onto the Agency's network servers.

When creating bookmarks and hyperlinks, the magnification setting *Inherit Zoom* should be used so that the destination page displays at the same magnification level that the reviewer is using for the rest of the document.

The bookmarks should be collapsed when document is opened so that all bookmarks are at the first level. Also see Section 1.1.13 for Initial View Settings.

1.1.12   Page Numbering

Only the internal page numbers of the document are expected (1-n). No additional page/volume numbers running across documents are expected. It is easier to navigate through an electronic document if the page numbers for the document and the PDF file are the same. To accomplish this, the first page of the document should be numbered page 1, and all subsequent pages (including appendices and attachments) should be numbered consecutively with Arabic numerals. Roman numerals should not be used to number pages (e.g., title pages, tables of contents) and pages should not be left unnumbered (e.g., title page.) Numbering in this manner keeps the Acrobat numbering in synchrony with the internal document page numbers.

The only exception should be where a document is split because of its size (Refer to Section 1.1.2 for information regarding File Size); the second or subsequent file should be numbered consecutively to that of the first or preceding file.

1.1.13 Initial View Settings

The initial view of the PDF files should be set as *Bookmarks* and *Page*. If there are no bookmarks, the initial view as *Page* only should be set. The *Magnification* and *Page Layout* should be set as default.

1.1.14 Security

No security settings or password protection for PDF files should be included. Security fields should be set to allow printing, changes to the document, selecting text and graphics, and adding or changing notes and form fields. The exception to this rule includes regulatory forms with pre-existing security and literature references that need to be copyright protected. At a minimum the receiver should be able to easily open and view the content.

1.1.15 Use of Acrobat Plug-Ins

It is appropriate to use plug-ins to assist in the creation of a submission. However, the review of the submission should not call for the use of any plug-ins in addition to those provided with Adobe Acrobat because regulatory authorities will not necessarily have access to the additional plug-in functionality.

1.2 XML Files

A working group at the World Wide Web Consortium (W3C) developed XML. It is a nonproprietary language developed to improve on previous markup languages including standard generalized markup language (SGML) and hypertext markup language (HTML).

XML is currently used for some content of the eCTD. The applicant should contact the applicant's own regional regulatory authority, understanding that other regulatory authorities may not accept these XML files.

Additional information about the XML standard can be found at the W3C Web site.

1.3 SVG Files

SVG is a language for describing two-dimensional graphics in XML. SVG allows for three types of graphic objects: vector graphic shapes (e.g., paths consisting of straight lines and curves), images, and text. Graphical objects can be grouped, styled, transformed and composited into previously rendered objects. Text can be in any XML namespace suitable to the application, which enhances searchability and accessibility of the SVG graphics. The feature set includes nested transformations, clipping paths, alpha masks, filter effects, template objects, and extensibility.

SVG drawings can be dynamic and interactive. The Document Object Model (DOM) for SVG, which includes the full XML DOM, allows for straightforward and efficient vector graphics animation via scripting. A rich set of event handlers such as onmouseover and onclick can be assigned to any SVG graphical object. Because of its compatibility and leveraging of other Web standards, features like scripting can be done on SVG elements and other XML elements from different namespaces simultaneously within the same Web page.

The specific use of SVG in a submission should be discussed with the regulatory authority.

Additional information about the SVG specification can be found at the W3C Web site.

1.4　　　Study Dataset Files

Specific regions include study datasets and may have different rules regarding the following topics:

- Allowable file formats

- Dataset file sizes

- Dataset filenames and allowable characters

Refer to Regional/Module 1 Implementation packages for additional information.