Quality eCTD Submissions

For successful transition from paper to e-submissions, the scientist and the information systems professional need to understand each other's tasks and responsibilities.

SIDNEY B. RUBINSTEIN, B.S., R.PH. AND JAMES E. CARTER, PH.D.
Regulatory Compliance Initiatives, Inc.

JEFFERY L. CARTER, B.A., M.S.
HighTech-Solutions, Inc.

PAPER REGULATORY SUBMISSIONS are becoming historical documents—e-submissions are the future. The basic principles for a successful transition are:
• Early planning and preparation;
• Knowing the regulatory science;
• Understanding the guidance documents;
• Understanding the ICH CTD format and FDA content specifications;
• Watching consistencies successfully;
• Understanding XML (eXtensible Markup Language);
• Knowing the e-submission process and the electronic backbone;

QUALITY eCTD SUBMISSIONS

• Paying attention to lessons learned; and
• Purchasing the right tools.

Riding this transition successfully can save your organization money, increase submission accuracy, decrease review times and increase the ability of both your company and the FDA to access files already submitted. Those who don't transition effectively will lose a competitive advantage. This article provides principles for developing a quality paper submission, followed by easy to understand instruction in developing a quality e-submission.

Early Planning and Preparation

Quality submissions begin with careful planning and preparation long before the IND, NDA, ANDA or DMF is assembled. Paper and electronic submissions are most successful when scientists consider regulatory guidances and the consequences of their actions before developing documentation of requirements such as the manufacturing processes, analytical methods validations and stability data. Methods and methods validation reports are successfully and efficiently assembled when developed with early attention to fonts, margins, abbreviations, terms related to the submission and consistencies across departments.

With proper planning and preparation, companies can have a clear vision of a quality eCTD submission long before they put pen to paper or fingers to keyboard. Whether the final submission is paper or electronic, the quality of the submission is built-in long before the data and information are gathered, assembled and submitted to the agency.

Know the Regulatory Science

The second principle of a quality submission is knowledge—knowledge of your molecule, the formulation, manufacturing process, analytical methods and specifications, as well as a thousand other details. This is what it takes to be good scientists. We should know our molecule and all the ramifications of our submission far better than any regulatory reviewer. However, knowledge is not enough. It is important to explain the issues and principles to the reviewer in clear language that is easily understood. Some individuals call this writing from the reviewer's perspective.

It is vital to understand what the reviewer needs to see and how the reviewer will view the information. In addition, it is important to be consistent. If we use the term "registration batches" in one section of the submission and then "stability batches" in another, the reviewer may be confused and will end up asking needless questions. The reviewer is as busy as you are and will put down the submission to attend meetings, accept phone calls and deal with interruptions. In essence, key information needs to be consistent and repeated to assure continuity in the review process without making the reviewer backtrack and waste valuable time. As we put togeth-

er a quality eCTD submission, we start with good science and knowledge of the reviewer's needs.

Understand the Guidance Documents

In large part, knowing the needs of the reviewer is knowing and understanding the guidance documents. FDA and International Conference on Harmonization (ICH) regulatory scientists provide us with valuable insights into their needs. The guidances are put together by knowledgeable scientists from regulatory agencies, academia and industry throughout the world.

The science associated with the submission should not be knowledge gained or presented in a vacuum. Knowledge must be related to the regulatory guidances. Key points from the guidances related to the submission must be communicated to all individuals contributing to the submission. It may appear that there are a hundred guidances with a thousand details, but in reality we digest this elephant one bite at a time.

For example, when developing active pharmaceutical ingredient (API) or drug substance data, be aware of the ICH guidance "Impurities in New Drug Substances"[1] and the allowable levels for each class of impurities. Note that companies current with the guidance are using unidentified impurities, identified impurities, specified unidentified impurities and total impurities. Unidentified impurities are generally given a limit of ≤0.1%, but this limit may be ≤0.05% if the total API dose is ≥2 g/day. Additionally, it is noteworthy that the term "Related Substances" is a historical term that is not the current language of the guidance or a quality submission.

Another detail is in the residual solvent specification. If in developing a molecule we are unaware of the ICH solvent classes and limits, we may create a synthetic manufacturing process that cannot be commercialized. If in setting the residual solvent specifications we only use ICH concentration limits based on inherent toxicity of the solvent, it is obvious to the reviewer that we do not understand the guidance. The specification is to be set within the ICH inherent toxicity limit, but also on the capabilities of the GMP process.

Similar details are present in the guidances for developing a quality submission for the drug product. The ICH guidance on "Impurities in New Drug Products"[2] and the ICH guidance on "Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances"[3] both indicate that "generally impurities present in a new drug substance need not be monitored or specified in the new drug product unless they are also degradation products."

Thus, it is important to distinguish and justify in a quality submission impurities that are process impurities from those that are degradation products. It is also important to distinguish potential from actual degradation products. The company must know the molecule and scientific information much better than the review-

er and must provide the reviewer with the information necessary to make an educated decision on approval of the submission.

Understand Format and Content Specs

Getting the CTD outline correct should be relatively simple. The key guidances related to the outline were first issued by ICH in 2001 and are listed as "M" guidances when they are multidisciplinary and "Q" guidances when they are related to quality or Chemistry Manufacturing and Controls (CMC) information. All guidances are available on the FDA website at http://www.fda.gov/cder/guidance/index.htm. Be sure to check both ICH guidances and ICH draft guidances. *(For the purposes of this article, the key format guidances for CMC information are listed as references 4-7.)* It is valuable that ICH recently responded to location questions by providing the June 2004 guidance: "Guidance for Industry, M4: The CTD—Quality, Questions and Answers/Location Issues."

For an IND, NDA or ANDA, the CMC section is the Module 3 outline. For the DMF format, companies are using the Module 3 Drug Substance section. Module 2 includes a summary Quality section and must not contain new information. The Quality summary is required for an IND, NDA or ANDA but is not required for a DMF submitted to FDA. Individuals writing a CMC Quality section should also be aware of administrative documents that are placed in Module 1 of the submission.

Although the groups comprising ICH could not agree on content for a CTD submission, they rightfully agreed on format. Subsequent to the agreements on format, FDA published draft guidances dealing with content for the drug product section and drug substance section (references 8 and 9 respectively). These guidances provide valuable information on the content for each section. For example, if the analyst performs stress degradation studies on the drug substance, is it part of analytical validation or is it part of stability? The guidance indicates it should be part of

---

Appendix 1: HTML vs. XML

Here are examples of how the codes differ.

HTML:

```
<HTML>
<HEAD>
<TITLE>Contacts</TITLE>
</HEAD>
<BODY>

<UL>
<LI>James E. Carter</LI>
<LI>jcarter@regulatorycomp.com</LI>
<LI>702 914 0798</LI>
</UL>
<UL>
<LI>Sidney B. Rubinstein</LI>
<LI>srubinstein@regulatorycomp.com</LI>
<LI>914 576 6412</LI>
</UL>
<UL>
<LI>Jeffery L. Carter</LI>
<LI>jeff@regulatorycomp.com</LI>
<LI>719 930 4407</LI>
</UL>

</BODY>
</HTML>
```

XML:

```
<?xml version = "1.0" encoding="UTF-8" standalone = "yes"?>
<DOCUMENT>
<CONTACT>
<NAME>James E. Carter</NAME>
<EMAIL>jcarter@regulatorycomp.com</EMAIL>
<PHONE>702 914 0798</PHONE>
</CONTACT>
<CONTACT>
<NAME> Sidney B. Rubinstein </NAME>
<EMAIL> srubinstein@regulatorycomp.com </EMAIL>
<PHONE>914 576 6412</PHONE>
</CONTACT>

<CONTACT>
<NAME>Jeffery L. Carter</NAME>
<EMAIL>jeff@regulatorycomp.com</EMAIL>
<PHONE>719 930 4407</PHONE>
</CONTACT>
</DOCUMENT>
```

Section 3.2.S.7.3 Stability Data. The data can support analytical validations, information on impurities, as well as the stability summary and conclusions. As an additional insight into knowing the guidances, both discourage use of the term "conforms" when reporting analytical test results. FDA prefers the use of descriptive terms or actual numbers. Understanding the guidances and putting together a quality submission applies equally to the electronic side of the submission.

Introduction to eCTD

Reasons for transitioning to an e-submission are basic: improve the submission and review process, increase accuracy of the submission and decrease total costs. All parties involved in the e-submission process share this common goal. Pharmaceutical companies transitioning from paper to e-submission must be careful to contain costs. Improving the process and accuracy must be translated into concrete improvements such as the ability to track and store data easier, a faster submission process, better products and increased personnel efficiency.

The "garbage in, garbage out" adage is very applicable to e-submissions. Not following guidances, not following the CTD outline and not watching consistencies will put garbage into your new XML e-submission process and will result in a garbage submission to FDA. Introduction of a new system, process and technology, together with a "garbage in" process, could introduce chaos to your submission process. The interaction between the scientists and the information systems group in your organization is critical. Getting the paper system in order is the first step toward a good e-submission.

Understanding XML

XML is a specification or standard that is used in eCTD submissions. XML enables an information provider (a regulatory submission from industry) and an information user (the regulatory authority) to create and exchange information. The content of information expressed in a mark up language is often referred to as "meta data." Meta data provides fundamental information about the information being exchanged. Mark up languages or meta data are typically used for three purposes: formatting, structuring data and data transport. Meta data can be visualized as a wrapper for information that enhances the appearance, increases understanding of the information, or enables the correct alignment or transport of the information to another source. *Appendix 1* contains examples of two different mark up languages, HTML and XML and the differences between the two.

Presently, the most popular mark up language is Hyper Text Markup Language (HTML). HTML comprises the meta data about a document that allows an internet browser to view the document. Speaking hypothetically, assume you need to gather information

Widely Used Acronyms

ANDA: Abbreviated New Drug Application
API: Active Pharmaceutical Ingredient
CERN: Centre European pour la Recherche Nuclaire
CMC: Chemistry, Manufacturing and Controls
CTD: Common Technical Document
DMF: Drug Master File
DTD: Document Type Definition
eCTD: Electronic Common Technical Document
FDA: Food and Drug Administration
GML: General Mark up Language
GMP: Good Manufacturing Practices
HTML: HyperText Markup Language
ICH: International Conference on Harmonization
IND: Investigational New Drug
NDA: New Drug Application
PDF: Portable Document Format
SGML: Standard Generalized Markup Language ISO 8879:1986
XML: eXtensible Markup Language

from a number of different distributed sources. To accomplish this task one might first define how the information authors should structure, format and deliver the information before sending it. XML is a tool that allows any user to define theses components. Document Type Definition (DTDs) and schemas are used to define your XML and are addressed later in this article.

There are a number of different mark up languages including: Standard Generalized Markup Language (SGML), HTML, XML and Generalized Markup Language (GML). XML evolved from SGML, and, in fact, many consider XML to be SGML version 2. Each markup language has its own strengths and weaknesses. HTML is great for formatting but weak in providing structure. XML is fairly good at formatting, excellent with structure and data transport, but is considered complex. Furthermore, XML is not as broadly accepted as HTML, which limits the number of authoring tools or viewing applications.

DTDs and Schemas

Somewhere along the transition path from paper to e-submissions you will hear about DTDs and schemas. As stated previously, these tools are used to define your custom XML. These document types are very similar. The purpose of DTDs or schemas is to let you know how to prepare your XML document for submission. These documents define how to communicate and process the data.

A good XML authoring tool will link your XML document to the DTD or schema, flagging potential errors in the markup. The syntax of DTDs and schemas are very different but they fulfill the same purpose. DTDs and schemas provide the framework. Writing about the differences is beyond the scope of this article. *(If you wish to understand the differences in greater detail, an excellent article can be found at http://www.xml.com/pub/a/1999/12/dtd/)*

ICH has designed the eCTD using XML and it defines it using DTDs and schemas that are available for download. These DTDs and schemas enable organizations to create XML documents needed for the e-submission. The FDA DTDs and schemas are the framework and protocols needed for the submission. Creating an e-submission, in effect, marries the FDA review process with the submitting organization, enabling the exchange of regulatory information. ICH, knowing such a relationship must work, has built numerous mutual benefits into their markup language, which include document IDs for easier tracking, checksum for document authentication and an open and flexible standard such as XML.

Before the e-Submission

Walk before you run. An e-submission to FDA does not require the purchase of hundreds of thousands of dollars in electronic tools. *(Some of the tools used at Regulatory Compliance Initiatives [RCI] are discussed later in this article.)* Tools are available to automate the FDA e-submission process and decrease the submission time through automation. These tools are fine but tend to keep the end user at a superficial level of understanding.

Having only a superficial understanding of XML, the FDA e-submission process and the electronic backbone can create problems. You need someone that intimately understands the electronic side of the submission. A fancy tool by itself will not be sufficient. Automated tools are only of value if they save you time and money. Here are a few parameters that should help in your decision making process.

Costs associated with automated tools include:
• Electronic equipment down time;
• Training;
• Equipment purchases;
• Maintenance; and
• Next-door help; i.e., loss of productivity for the person next door that is computer literate. Lost productivity can be as much as 15-30 percent.

Savings associated with automated tools:
• Faster submission time;
• Faster search and retrieval;
• Increased accuracy; and
• Improved product to the customer.

When purchasing an electronic tool, include the requirements of three participants in the process: scientist, information systems professional and FDA.

Depending on the company size, potential hidden costs could include increased disk space, a database, a hash calculator, Adobe Acrobat, an XML authoring tool and a word processor. Remember, walk before you run. It is not advised to jump straight into a high-dollar, fully automated e-submission tool. There are plenty of smaller, completely capable tools that will enable you to walk before sprinting into a fully automated and more expensive tool.

About the Authors

James E. Carter is a principal in the consulting firm Regulatory Compliance Initiatives, Inc., (RCI) and has more than 25 years experience in the pharmaceutical industry, including more than 11 years as a consultant. He specializes in preparation and filing of INDs, NDAs, ANDAs, and DMFs in electronic Common Technical Document (eCTD) format. Dr. Carter is an expert in ICH and FDA guidances and current Good Manufacturing Practice (cGMP) regulations.

Jeffery L. Carter has more than 12 years experience in implementing technology and information systems. A program manager for the Motorola Semiconductor Product Sector for more than five years, he designed, developed, and implemented an XML-based knowledge management system. As a telecommunications and open source expert, he is experienced with such technologies as intelligent agents, natural language processing, vector space analysis, ontologies, Oracle, CORBA, Perl, PHP, MySQL, Linux and XML.

Sidney B. Rubinstein is a principal in the consulting firm Regulatory Compliance Initiatives, Inc. (RCI) and has almost 30 years experience in the pharmaceutical industry, the last 20 years as a consultant. He specializes in current FDA and international cGMP regulations, and audits and prepares API and finished dosage manufacturers for FDA Pre-Approval Inspections (PAI). He also specializes in the preparation and submission of Drug Master Files in eCTD format.

The e-Submission

The e-submission process starts long before you request your submission number from FDA. As stated previously, your e-submission process starts with knowing and using the guidances, knowing the CTD outline, following the content for each section/document, and watching for inconsistencies. With your first e-submission, FDA will probably suggest a sample number for your submission. If FDA does not make this suggestion, make the suggestion; request a sample submission number for your first couple of submissions. This is an excellent opportunity to work out the kinks in your process/system and open the communication channels with FDA. The sample submission does not take that much extra effort, is an excellent opportunity and is worth the investment.

To complete an e-submission to FDA, RCI uses the following tools:
• XML authoring tool: XMLSPY (www.altova.com/)
• Checksum calculator: HashCalc (www.slavasoft.com/hashcalc/)
• Word Processor: Microsoft Word (www.microsoft.com/)
• PDF converter: Adobe Acrobat 6.0 Professional (www.adobe.com/)

In general, the RCI e-submission process is outlined as followed:
1. Assemble the backbone.
2. Scan the non-electronic material.
3. Convert and parse the submission into PDF documents and place them into the backbone.
4. After parsing and PDFing, build the XML document using XMLSPY.
(RCI leaves the document ID number and checksum calculation as the last step in the XML file creation.)
5. Ship the package—burn the CD and place the CD in a prepared folder with the hard copy cover letters.

Lessons learned at RCI include the following:
• FDA does not accept any PDF versions later than 1.4;
• The checksum in the XML needs to be lower case;
• Be sure to have a table of contents and bookmarks in your PDF;
• A tool that enables you to link your DTD or schema to your XML document increases accuracy;
• Extra node extensions are not allowed in XML; and
• Walk before you run.

Conclusion

For an CTD submission, it is imperative that the company works as

eCTD Top 10 List
Here is CDER's top 10 list of comments/mistakes in e-submissions:

10. Files referenced in the XML backbone(s): If the file is unreferenced, the reviewer will not be able to find the document.
9. eCTD submissions include module 1: Module 1 contains essential information for the submission. It is a must for the submission.
8. Application numbers are six digits: The application number is six digits in length, no spaces, letters, or special characters; just six digits.
7. Sequence numbers are four digits: The sequence number is four digits in length, no spaces, letters or special characters; just four digits.
6. Unneeded node extensions are removed: Node extensions are unwanted and unneeded.
5. MD5 checksum are correct: Checksum must be lower case.
4. Documents conform to eCTD granularity: Do not combine documents; keep the appropriate level of granularity.
3. XML must be standard components: Avoid custom DTDs and schema.
2. PDF hyperlinks/bookmarks are correct: Double check to make sure all hyperlinks and bookmarks are valid.
1. PDF documents include TOCs: a hyperlinked table of contents is required.

a team to develop and submit quality documents that are consistent with the guidances and internally consistent in terms. But the real advantage for a company and reviewer is to take the quality CTD submission and convert it to an eCTD document. The scientists and the information systems professionals need to increase their understanding about each other's needs in order to successfully complete an e-submission. Increasing understanding results in a smooth transition. If necessary, essential training should be obtained so that your organization can remain competitive.

A successful transition from paper to e-submissions has both the

scientist and the information systems professional taking one step toward understanding each other's tasks and responsibilities. The scientist must not only know the science, understand the guidances, the CTD format and content, and be sure to maintain consistencies, he must also take a step closer to understanding XML, know the e-submission process and understand the electronic backbone. The information systems professional must not only purchase the right tools for the job, understand XML and understand the electronic backbone, but must take one step closer to understanding the guidances, understanding the CTD format and content, and watching consistencies successfully. Riding this transition successfully can save your organization money, increase the accuracy of the submission and decrease review times, giving your company a competitive advantage. ◼

References

www.xml.org
www.w3.org
www.extropia.com
www.xml.com/pub/a/1999/12/dtd/
www.zakon.org/robert/internet/timeline/

1. "Q3A Impurities in New Drug Substances, Revision 1," February 2003.

2. "Q3B(R) Impurities in New Drug Products, Revision 1," November 2003.

3. "Q6A Specifications: Test Procedures and Acceptance Criteria for New Drug Substances and New Drug Products: Chemical Substances," December 2000.

4. "Submitting Marketing Applications According to the ICH-CTD Format — General Considerations," August 2001.

5. "M4: Organization of the CTD," August 2001.

6. "M4Q: The CTD — Quality," August 2001.

7. "M4: The CTD — Quality, Questions and Answers/Location Issues," June 2004.

8. "FDA Draft Guidance: Drug Product Chemistry, Manufacturing, and Controls Information," January 2003.

9. "FDA Draft Guidance: Drug Substance Chemistry, Manufacturing, and Controls Information," January 2004.

10. CDER eCTD Top 10 List.