

## The Untapped Potential of the Protocol Representation Model

Jeffrey Abolafia, Rho Inc., Chapel Hill NC  
Frank Dilorio, CodeCrafters, Inc., Philadelphia PA

### ABSTRACT

Recent FDA guidances have established CDISC models such as ADaM and SDTM as submission standards. As a result, most organizations have focused CDISC implementation strategies on SDTM and ADaM, which have usually led to higher costs and longer timelines. Moving standards implementation “upstream” can maximize the value obtained from CDISC standards. One largely overlooked standard is the Protocol Representation Model (PRM), the beginning “end” of “end-to-end.” The PRM has the content and potential to streamline research throughout the entire product life cycle.

This paper is an overview of the PRM. It describes what it is; discusses the business case for its implementation; describes Rho’s implementation strategy; demonstrates how its use at Rho has improved operations; and presents strategies for collecting and storing PRM metadata. The paper should give the reader an appreciation of the content and scope of the PRM and its use beyond simply storing protocol items as metadata.

### INTRODUCTION

One of the major goals of CDISC is to provide end-to-end standards. With the release of results level metadata in 2015, this objective has finally been realized. The addition of analysis results metadata not only provides standards from protocol to analysis results, but also full traceability from collected data to results.

A second goal of CDISC is to provide standards that improve the efficiency of clinical trial operations. However in order to gain the full value of CDISC standards, they must be implemented as far upstream in project workflow as possible. Currently, many organizations have implemented the SDTM and ADaM standards, a small percentage of organizations have adopted the CDASH data collection standard, but very few have taken advantage of the PRM. As a result, most organizations have not realized the full benefit of CDISC standards. In fact, by using a downstream implementation approach, most sponsors are actually spending *more* time and money preparing data for regulatory submission.

In the sections below we describe what is the “Protocol Representation Model” (PRM), what is contained in Version 1.0 of the PRM, the business case for its implementation, several use cases for PRM, how Rho, Inc. has utilized the PRM, and recommendations for implementation. The focus of the paper is on the content of PRM and not on the technical details of the underlying model that houses PRM concepts.

### BACKGROUND

A construction crew does not simply show up at a job site with materials and good intentions, hoping that they will somehow assemble a safe, functional structure. The blueprint, a result of planning, resource allocation, and other processes, guides the project from groundbreaking to ribbon-cutting. The blueprint is essential, and it is at the headwaters of the project’s workflow.

In a similar manner, the intricacies of clinical research also require a guiding document – the study protocol. This document describes the purpose of the study, the resources needed to conduct it, its design, a schedule of milestones, criteria for inclusion, evaluations to be performed, and a host of other items relevant to an organization’s managers, statisticians, compliance personnel, and others. The protocol is also essential for external users, notably regulatory agencies.

Rich in essential information, the protocol has typically been written and stored in electronic formats such as Microsoft Word. Although such a document conveys essential features such as study intent, design, and management, it is *not* possible to use it as a data source. The .DOC version of the protocol, for example, identifies inclusion and exclusion criteria, but it isn't possible to programmatically access the document and create metadata for the SDTM IE Domain. The criteria must instead be copied or reentered into a metadata database. This unnecessarily consumes resources and is error-prone. The study protocol was, in the early 2000's, a resource fairly crying out for standardization and programmatic re-use.

CDISC and leaders of HL7 and the FDA began a protocol standardization project in 2002. The goal was to identify and organize protocol elements that are common to most clinical studies. The International Conference on Harmonization (ICH) guidances E6, E3, and E9 and the requirements for registration of studies in EudraCT were used to identify information and concepts contained in clinical study protocols. Later, protocol elements were added to meet the requirement of WHO and clinicaltrials.gov.

Version 1.0 of the CDISC Protocol Representation Model (PRM) was released in 2010. The rationale for developing the PRM was summarized on the CDISC website (1) as:

"Protocol v1.0 was developed to support: a) protocol document generation; b) research study (clinical trial) registration and tracking; c) regulatory oversight and review; and d) single-sourced, downstream electronic consumption of protocol content, allowing users to create and quality control content once, and reuse for trial registries, protocol and case study report templates, SDTM study design and more"

Subsequent versions of the PRM have not been released. This is surprising, given that the PRM focuses on study startup, where the greatest value of implementing CDISC standards is obtained. What is also surprising is the relative dearth of papers at pharmaceutical industry conferences that touch, much less focus on, the PRM (see "Recommended Reading," below). Clearly, the PRM has been lightly embraced by its intended audience. Even though it is one of CDISC's "foundational" models, in the same grouping as the more familiar CDASH, SDTM, and ADaM, it is conceptually different. Just what it is, why it is different, and how it can be used are addressed in the next section.

## WHAT IS THE PRM?

Many CDISC models – notably CDASH, SDTM, and ADaM – have become part of the working vocabulary of pharmaceutical industry workers and their regulatory counterparts. Likewise, the means for describing the data – define-xml – has evolved from being a new technology to one that is commonplace. Since the PRM is a new way to represent part of an electronic submission, it should just be more of the same, right? Just read the CDISC implementation guide, then deliver the protocol in XML.

Actually, it's not quite that simple.

The PRM is simply a *conceptual* model for organizing a protocol. The model identifies over 350 items typically found in a protocol and organizes them into a common structure intended to be machine-readable. As noted above, study documents stored in Word or similar formats cannot be programmatically accessed for downstream usage. If Inclusion/Exclusion criteria were stored in a predictable, machine-readable format, it would be a trivial task to extract the values and create the SDTM TI domain. The PRM's identification of programmatically accessible protocol elements enables domain creation to be automated. It can also automate a host of other previously manual activities.

There are several key points to understand when considering how to use the PRM:

- There is no "PRM deliverable." Unlike SDTM, ADaM, SEND, define-xml, the PRM does not describe a submission deliverable. Rather, it is a *model* for representing how the protocol can be organized for electronic access during the course of a study. The PRM describes how protocol elements are organized and what they contain. Not all the model's elements need to be populated, but those that are have content that is precisely defined, predictable and programmatically accessible.
- You don't have to model PRM elements using the Unified Modeling Language (UML). UML was initially chosen because the CDISC BRIDG model (which includes PRM) used it and because it was a

robust, visual mechanism to identify entities and hierarchical relationships. UML could also provide semantic consistency and interoperability with other BRIDG standards. The PRM document (see "Appendix A: Resources - CDISC") contains nearly 200 pages of detailed documentation of UML classes, attributes, and the like. Even for the "initiated," much less those unfamiliar with UML, the sheer bulk and complexity is more than a little intimidating. If an organization's software development team is more familiar with another data representation paradigm, there's no reason why it cannot be used.

- Since it is not a deliverable in the traditional sense (not like ADaM or SDTM datasets, for example), it can be stored in a format most comfortable to the developer/user. Rho chose Oracle in part because all of our other CDISC-related metadata is held in Oracle tables. The Oracle environment is secure, robust, and familiar, thus making it an easy choice (Rho's implementation is discussed in detail in "Use Cases," below).
- Current and future versions of the PRM are and never will be exhaustive. After reviewing the contents of the PRM model it is not uncommon to find missing elements. Because the PRM is extensible, all that is required is that the extensions to the model be fully described in the database schema and, if the protocol will be shared among organizations, in an ODM extension (described more fully in "Data Exchange," below).
- PRM implementation means protocol components are stored in a database and, therefore, are programmatically accessible to anyone familiar with the database design. This has profound implications for automation of work processes: elements can be used in multiple studies; cut and paste from Word or Excel documents becomes a thing of the past. If our experience with SDTM and ADaM metadata is any indication, the range of PRM-based applications will only grow as users gain familiarity with the model. This has nothing but positive implications for submission quality and delivery timelines.

## IMPLEMENTATION

Metadata forms the foundation for ADaM and SDTM datasets and documentation. Machine-readable dataset and variable attributes can be accessed "end-to-end," throughout the study life cycle. The metadata and tools to make it accessible to programs drive workflow from data capture to define-xml. PRM tables and elements, regardless of database choice and design, play the same role: they facilitate writing the protocol and automate previously manual tasks.

For the sake of discussion here, we have broken down PRM implementation into several highly generalized pieces: database, interface, access, and data exchange.

**Database:** Some database issues have already been addressed: data modeling and storage are the choice of those implementing the PRM. Key features here are security, audit trail implementation, ease of access by applications using the PRM data, extensibility, and scalability.

**Interface:** Creating a study protocol consistent with PRM or similar models is at its core a form of structured authoring. Ideally, the interface, whether it's a commercial product or home-grown, should follow the flow of the process it's supporting. It should also be able to make similar content from other protocols accessible. In a group of related studies, for example, the interface should be able to let the user easily copy inclusion/exclusion criteria from an existing study's protocol to one being developed. Thus the interface is both a means of entry of new metadata as well as accessing a repository of existing materials. Screen shots of the system developed at Rho are found in "PRM Implementation at Rho," below.

**Access:** Even the best-designed PRM-compliant database will be next to useless if its contents are not easily accessible. The "Interface" section, above, alluded to this when it described building a protocol with items taken from other protocols. This required transparent access to a protocol repository. Without a repository, and without an interface to easily access needed items, efficiencies in protocol development would be dramatically reduced.

Once the electronically readable protocol is developed, many types of applications present themselves, ones that were not possible when the protocol was cocooned in a Word(-like) format. This is where

protocol access tools come into play. Software (e.g., SAS macros) that has well defined inputs/outputs, is easily used, and that can access the protocol database can be used throughout the study. Applications include: creating CDASH-compliant CRFs; seeding SDTM tables; and providing content for the study Reviewer's Guide; and supplying content for top-level elements in define-xml.

**Data Exchange:** As mentioned earlier, the PRM is an extensible model for protocol description. The database tables storing the protocol metadata are not deliverables. Sometimes, however, all or part of the study protocol metadata may need to be exchanged. The CDISC Study Design Model (SDM) is one delivery mechanism (2). It is a schema extension of ODM 1.3, and describes Study Design features of a protocol (experimental design, schedule of activities, eligibility criteria, and summary information). The advantage to using SDM is familiarity. Just as database tables are "massaged" when creating an ODM-compliant CRT-DDS (define-xml) file, so too would PRM tables be transformed to fit into the SDM's structure. SDM extensions can be written for elements that were defined by the PRM but not in the current SDM schema.

## THE BUSINESS CASE FOR PRM

As noted earlier, CDISC data models are the *de facto* data standards for submitting regulatory data. Indeed, they are required for all studies starting in December 2016. As a result, the data standards implementation strategy deployed by most organizations is simple: "get the FDA what they want". To that end, standards implementation has been focused on SDTM, ADaM, and define.xml. This "downstream deliverables" strategy has created additional deliverables and increased costs (3).

Moving standards implementation further upstream, however, can provide much greater value to an organization. A business case study on CDISC standards by Gartner (4) found that implementing standards from the beginning can save up to 60% of non-subject participation time and cost. About half of the value was gained in the startup stages. The study also reported that the average study startup time can be reduced from around five months to three months. A second study by Medidata (5) looked specifically at PRM. It reported that storing protocol concepts as structured data has significantly reduced the number of protocol amendments, the recruitment cycle, the number of handoffs, the time for protocol review, and the overall time for protocol development.

One obvious reason to adopt the PRM is to streamline protocol development. The PRM standardizes over 350 concepts usually found in clinical protocols. When these standardized concepts are stored in a database, they can be inserted into protocol templates using structured authoring tools. They can also be re-used for subsequent protocols. This creates semantic consistency across protocols, makes it easier to understand protocols, and facilitates searches for key information buried in a protocol. This benefits sponsors and regulatory reviewers alike. Submitting protocols in a standard and consistent format to regulatory reviewers can decrease the amount of time for protocol review, facilitate discussion in Pre-IND meetings, and accelerate product development.

The PRM not only has the potential to streamline protocol development but can also facilitate creating databases and documents further downstream. The PRM's study design data can be exported to Clinical Data Management Systems (CDMS). Over 20 elements of the PRM map to elements in SDTM trial design datasets and to elements required for submitting to ClinicalTrials.gov. The PRM can also assist in preparing the following regulatory documents:

- PIND/PIDE meeting package
- INDs and CTAs
- Annual Reports/DSURs/PSURs
- EOP2 meeting package
- Pre-NDA/BLA meeting package
- NDA/BLA/Marketing Authorisation Application
- 120-Day Safety Updates

Several of these use cases are presented later in the paper.

## PRM IMPLEMENTATION AT RHO

Rho, Inc. has a long and successful history of implementing metadata-based systems. In the early 2000's, we created Microsoft Access databases that held metadata contributing to the creation of define.pdf and XPT files supporting electronic submissions. We later added metadata for the creation of Tables, Figures, and Listings. During this period, and later, for SDTM, ADaM and define.xml, we realized that metadata without tools that made it accessible was marginally helpful at best. We developed a metadata entry interface and created many tools that made access to the now Oracle-based data as transparent as possible. Development cost was non-trivial, but the payback in client satisfaction, reduced time of deliverables, and (not to be discounted!) the "cool factor" more than justified the investment.

So when, at the direction of Rho senior management, a PRM implementation project was proposed, a proprietary, "home grown" system was a natural choice. There were several reasons for undertaking this initiative:

1. to follow the end-to-end strategic plan for standards implementation outlined by the authors in 2012 (3)
2. our success with dataset, analysis results, and study level metadata in streamlining processes and clinical trial operations, as noted above
3. having a single place where information about a trial is entered and is then re-used throughout a study or the entire product development life cycle.

A project team was assembled with representation from regulatory, project management, data management, biostatistics, clinical operations, and statistical programming departments. The diversity of the group – medical writers, programmers, regulatory affairs, and others – speaks to the range of uses of the PRM's contents.

Based on our previous experience with metadata design and implementation, we knew metadata "mission creep" was positive and to be encouraged. Focusing only on making protocol development more efficient would have ignored other, downstream uses of the metadata. The team guided development of the database and interface. Select "priority elements" were implemented first: high-level, descriptive elements, trial design, and schedule of events.

After the initial release of the system, a slightly differently configured team has remained in place to identify tasks on the "what's next?" list. The tasks are discussed at the end of the next section.

We felt implementing the PRM metadata model as far upstream as possible would provide maximum value and the greatest opportunity for re-use. Therefore, it made the most sense to begin with the protocol. The PRM served as our starting point. If a concept or controlled terminology existed in PRM, we used it. We then expanded the PRM to include additional concepts at both the study and program level (for example, in addition to number of subjects *enrolled* we also wanted to track the number *screened*. We also added elements that would enable us to populate the SDTM TS domain; this is described in more detail in "Using PRM Data: SDTM Trial Design Datasets," below).

As we have already noted, and as we will see in the discussion below, a key aspect of the PRM is that it can be used not just for assembling a protocol but for other study parts as well. Rho's experience with submission ("eSub") metadata demonstrated the value of extending and reusing it beyond its purported use. This is illustrated for both the PRM and eSub databases in **Figure 1**, below.

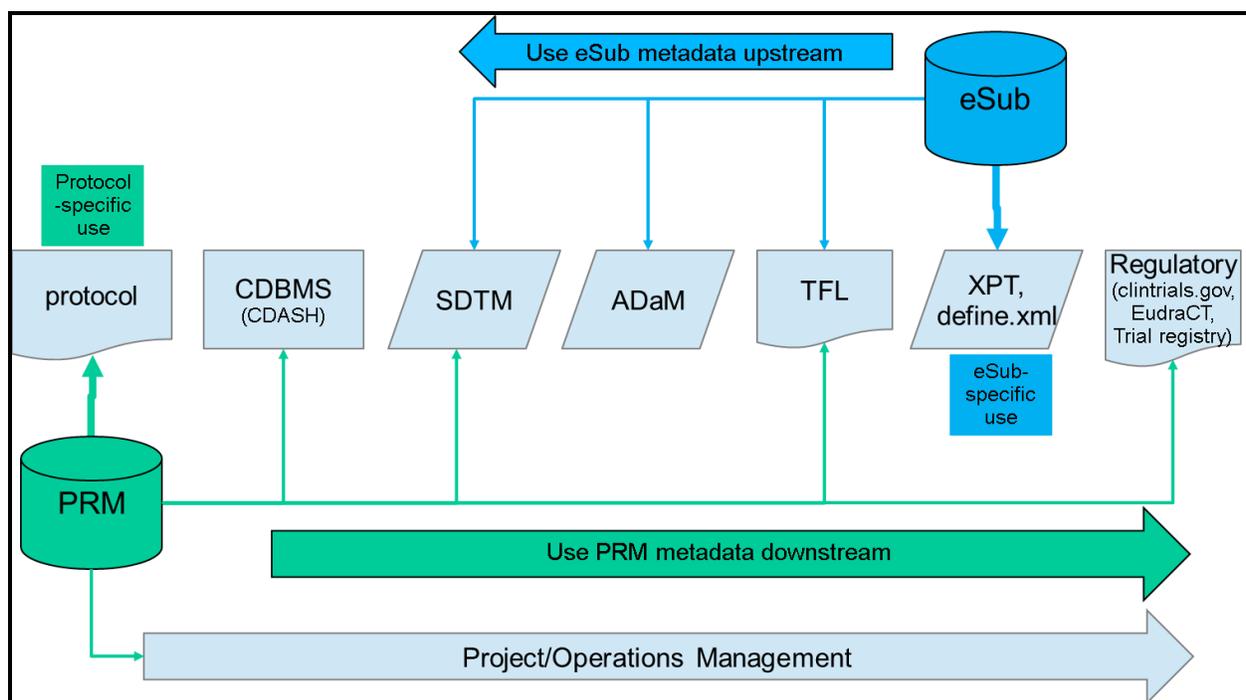


Figure 1: PRM, eSub Metadata Re-Use

## USING PRM DATA

In this section we describe how Rho is using the metadata stored in our PRM. Some of these use cases are in production while others are still in progress.

### [1] PROTOCOL DEVELOPMENT

Not surprisingly our first use case is protocol development. Protocols are currently developed as Rho, and most of the clinical research industry, as Word documents. We decided that moving directly to a “structured authoring approach” would require too large of a culture change. As a result, we decided on a more conservative approach. We started by adopting the TransCelerate protocol template (6) as our Rho Standard. We then mapped PRM concepts to the TransCelerate protocol template. Once protocols were developed in Word, we entered protocol related information into our PRM. After this was completed, we had a database with machine readable protocol metadata that could be re-purposed downstream. In the long term, we plan to start a study by first entering protocol information into our PRM database and then use structured authoring tools to read the PRM data and populate our standard protocol template.

### [2] CLINICAL DATA MANAGEMENT SYSTEM SETUP

Adhering to our strategy of leveraging end-to-end standards, we are aligning our protocol with our clinical data management system (CDMS). Our PRM implementation includes a schedule of events module, which consists of a matrix of Case Reports Forms (CRFs) and the visits at which they are collected. The schedule of events and the study design can be exported into the CDMS. This not only facilitates setting up the CDMS for a study, but also ensures that the protocol and what is collected in the CDMS are aligned.

### [3] SDTM TRIAL DESIGN DATASETS

At this point in time our PRM has had the most immediate impact on streamlining the production of SDTM trial design datasets. Over 25 concepts in our PRM map directly to the SDTM Trial Summary (TS) dataset. In addition, we have harmonized the controlled terminology in our PRM with the CDISC PRM. Next we did a gap analysis to find concepts needed for TS that are missing from the PRM. These

concepts are currently being added to our PRM. Once this task is completed, we expect to be able to create the TS datasets directly from PRM with little or no additional programming.

The Trial Inclusion/Exclusion (TI) trial design dataset served as the easiest SDTM dataset to create directly from our PRM. **Figure 2**, below, shows the interface to the Oracle database for the TI elements:

Design	Agents	Objectives	Endpoints	Intervention	Observation	Population	Subject Selection	
<b>Inclusion Selection Criteria</b>							<b>Sequence</b>	
Age 12 months to less than 48 months, either gender.							1	
Clinical history of peanut allergy or avoidance of peanut without ever having eaten peanut.							2	
Serum IgE to peanut of > 5 kUA/L determined by UniCAPTM.							3	
Wheal = 3mm on skin prick test to peanut extract compared to a negative control.							4	
A clinical reaction as defined in Protocol Section 6.4.1.3 at or below ingestion of 1 g peanut flour (500 mg peanut protein) during screening blinded OFC.							5	
Written informed consent from parent/guardian.							6	
<b>Exclusion Selection Criteria</b>							<b>Sequence</b>	
History of severe anaphylaxis with hypotension to peanut.							7	
Documented clinical history of allergy to oat.							8	
Suspected allergy to oat and a wheal greater than or equal to 7mm on skin prick test to oat extract compared to a negative control.							9	
Chronic disease other than asthma, atopic dermatitis, rhinitis requiring therapy; e.g., heart disease or diabetes.							10	
Active eosinophilic gastrointestinal disease in the past 2 years.							11	
Participation in any interventional study for the treatment of food allergy in the 6 months							12	

**Figure 2: User Interface for SDTM TI Elements**

The code below can be used for any study to easily create the TI dataset:

```
%setup(program=T:\Submissions\Rho\CDASHToSDTM, study=CDASHToSDTM01)

data ti;
  set ora.v_study_incl(in=ini)
    ora.v_study_excl(in=ine) ;
  where study_uid = &studyid ;
  domain = "TI" ;
  if      (ini) then iecat = 'Inclusion' ;
  else if (ine) then iecat = 'Exclusion' ;
run;
```

In the example, the setup macro allocates the necessary libraries and returns study specific information to the program. We have developed similar modules/program to automate creating the Trial Arms (TA), Trial Visits (TV), and Trial Elements (TE) SDTM data sets. Two additional benefits derived from using this method to create trial design datasets are: 1) the information is entered by research assistants not programmers. This not only saves money, but also allows programmers to use their time more effectively; and 2) all of the information in trial design datasets is in synch with the protocol – if the protocol is amended, the information is changed in only one place, the PRM database.

A not-unexpected reaction by any programmer reading the code snippet on the previous page would be “generalize it.” As was the case while developing our eSub tools, so it is the case here: tasks that are able to be generalized and will be repeatable are defined as SAS macros. Similar uses of the metadata can be described and implemented as a macro library. This makes access to the database possible with a minimum of programming effort.

#### [4] OPERATIONAL AND STATISTICAL REPORTING

In the past year, Rho has developed an interactive application to produce operational and statistical reports. The application contains a library of report templates that can be run as is or customized by the end-user. The application is also linked to our PRM, allowing it to automatically be aware of study-specific information such as title, sponsor, study design, treatment groups, and the schedule of events. As a result, most reports in the system can be automatically produced by end-users, including those with no programming background.

#### [5] MANAGEMENT/TRACKING

One of the many benefits of any metadata-based system is that it is inherently multi-use. A metadata source table can easily be used for multiple types of tasks throughout the project life cycle. While our PRM was designed to improve trial operations, we are also now using PRM metadata as a management tracking system for all of our projects. Our PRM has provided us with a machine readable database that contains detailed information about each protocol. This includes: title, sponsor, phase, study design, blinding, therapeutic areas, duration, planned number of subjects and sites, study start dates, study status, and other PRM fields. Using this database as input, one can easily determine the number of active studies, produce a listing of all oncology studies, or display all pain studies using a parallel group design. **Figure 3**, below, displays a subset of field and records in our PRM database.

StudyUID	ShortTitle	Phase	StudyType	Configuration	PlannedNumberSubjects	Multicenter	IdentificationNumber	StudyBrandNameLong	StudyBrandNameShort	PlannedNumberSites	Duration	Title	TrialPurposeClassification
14	ARA08	II	ND Exempt Intervention	Parallel Group Design	200	Y	ARA08	StopRA		21	3 YEARS	Strategy to Prevent the Onset of Clinically Apparent Rheumatoid Arthritis	Safety and Efficacy
8601	Prevent CMV	NA	Observational		80	Y	CTOT-22			5	18 MONTHS	Prospective Multicenter Cytomegalovirus (CMV) Specific Immune Monitoring to Stratify Patient Risk and Guide Antiviral Prophylaxis after Lung Transplantation (PREVENT- CMV)	Proof of efficacy
37	Interferon		Mechanistic		120	N	ADRN-01		Interferon	1	5 YEARS	Investigation of Reduced (Study Drug) Responses in Peripheral Blood Mononuclear Cells of Participants with Atopic Dermatitis and a History of Eczema Herpeticum	

**FIGURE 3: THE PRM DATABASE**

## WHAT'S NEXT

In the next year we plan on enhancing the PRM capabilities described above and rolling out several new initiatives. Rho works on both commercial drug development projects and on federally sponsored projects. As a result, we are required to report to several registries. These include ClinicalTrials.gov, EudraCT, Immport, and TrialShare. Many (over 25) of the concepts contained in PRM already map to ClinicalTrials.gov. We are in the process of mapping additional concepts in our PRM to ClinicalTrials.gov and then to the other registries listed above to streamline reporting requirements.

A second high priority initiative is linking our PRM to our other data systems. Our PRM already contains a module where one can specify the data systems (i.e. EDC, IVRS, Safety) needed for a project. When setting up a project in PRM, the end users will indicate the data systems required for the project. For Phase One of this initiative, our PRM will send an alert to a group whose services are required for the project. For example, if EDC is selected in PRM, an alert will be sent to data management informing them that there is a new study that requires an EDC component. For Phase Two, output from PRM will be used for setting up data systems for new projects.

An item given higher priority than initially planned came from our regulatory and medical writers. As mentioned in the "Protocol Development" section, moving to a full-fledged structured authoring system

was seen as too great a paradigm shift from the way protocols are typically developed. Our use of a protocol template was "structured light," a first step toward the structured authoring goal. But the writers were used to the Microsoft Word environment and wanted to remain in it if possible. This led to development of an app that interacts with Word, embedding tags in the Word protocol template that can be used to extract protocol elements from the Word document and populate the protocol database. Conversely, protocol elements can be extracted from the database and used to populate the Word version of the protocol. The Word-based document becomes the interface to the metadata, and offers an alternative method of entry and retrieval.

It is worth noting that there are many more use cases for the PRM that can possibly be discussed here. The downstream benefits of PRM are only limited by resources, priorities and our imagination.

## CONCLUSION

The PRM, the CDISC standard with the lowest adoption rate, has the potential to provide sponsors with the greatest value from implementing CDISC standards. In the past year, Rho has created a proprietary version of the PRM standard, storing protocol related concepts in a machine-readable format and developing an interface that allows information to be easily entered in a standard format by an end-user. This metadata, created during the startup phase of a project, can then be utilized throughout the life cycle of a project, thereby increasing efficiency and decreasing the amount of time and resources needed for project deliverables. At this point in time we are re-using PRM metadata to streamline: 1) setting up our CDMS and to align our CDMS with the protocol; 2) creating SDTM trial design domains; 3) operational and statistical reporting; and 4) managing and tracking for all of our clinical studies.

While we are only in the early stages of our PRM initiative, the initial results are promising. Having a single source for protocol related concepts has not only improved efficiency but has also led to higher quality and increased consistency throughout the life cycle of a project. In the next year we plan on extending our use of PRM to actual protocol development, reporting to registries such as ClinicalTrials.gov, linking PRM to inform our other data systems, producing SDTM and analysis datasets, and streamlining the creation of other regulatory documents.

An earlier paper by the authors (see "References," below) stated that well-constructed metadata and metadata access tools can have a significant, positive impact on the creation of the datasets and documents that comprise product development. The same is true for metadata describing a protocol: metadata-driven applications and utilities can be integrated into standard business processes, speeding the production and improving the quality of deliverables.

## REFERENCES

1. "Protocol Representation Model (Protocol) v1.0". Online at <http://www.cdisc.org/standards/foundational/protocol>
2. "Study/Trial Design Model". Online at <http://www.cdisc.org/standards/foundational/sdm-xml>
3. Abolafia, Jeff and Frank Dilorio, "Cost Effective Standards Implementation: A New Paradigm for the Drug Development Life Cycle". 2012. Proceedings of the PhUSE Annual conference.
4. "Gartner Report: CDISC Standards Enable Reuse Without Rework". Online at <http://www.cdisc.org/members-only/business-case-use-cdisc-standards>
5. Pines, Joshua and Michelle Marlborough, "Protocol Representation in the Real World". 2012. Proceedings of the PhUSE Annual conference.
6. "Common Protocol Template". Online at <http://www.transceleratebiopharmainc.com/initiatives/common-protocol-template>

## ACKNOWLEDGMENTS

We would like to thank our colleague Elizabeth Goodman for her valuable input during preparation of this paper.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies.

## RECOMMENDED READING

### From CDISC

As is the case with learning about any CDISC model, the best place to start is the CDISC web site: <http://www.cdisc.org/standards/foundational/protocol>.

The ZIP file downloaded by clicking [Download Protocol v1.0](#) contains these files:

1. PRM V1.0 Version Notes.pdf A manifest of the ZIP file's contents and suggestions for its use
2. Protocol Representation Model Document Version 1.0.pdf. The PRM reference, containing the history of PRM development, an overview of its contents, and the UML descriptions of classes and elements.
3. PRM V1.0 Diagram.xls The UML representation of the PRM, better viewed by the next item ...
4. BRIDG R3.0 with PRM V1 View.EAP The Enterprise Architect representation of the PRM; the machine-readable description of the PRM as described in item 2, above. Those interested in viewing the UML in EAP format can download a free viewer from Sparx Systems <http://www.sparxsystems.com/>

The web site also has a link to the Protocol Development Wizard. This tool guides the user through entry of some of the PRM elements held in an Excel file. The wizard was developed as a proof of concept implementation of the PRM. Its limited functionality and its lack of upgrading since its release several years ago do not make it a compelling choice for serious PRM use. Storage in Excel also mitigates against one of the most compelling reasons to use PRM (or any type of metadata, for that matter): reuse of XLS files across studies and projects, while not impossible, is certainly difficult.

CDISC Study Design Model in XML (SDM-XML)

[http://www.cdisc.org/system/files/all/standard\\_category/application/pdf/cdisc\\_sdm\\_xml\\_1.0.pdf](http://www.cdisc.org/system/files/all/standard_category/application/pdf/cdisc_sdm_xml_1.0.pdf)

### Other

Goud, Judith and Julie Smiley "Data Transparency Through Metadata Management"

<http://www.lexjansen.com/phuse/2014/pp/PP10.pdf>

Marlborough, Michelle and Joshua Pines "Real-World Application of the Protocol Representation Model"

<http://www.lexjansen.com/phuse/2012/cd/CD06.pdf>

Paletta, Deborah "Managing Data Standards Libraries Compliance"

[http://www.lexjansen.com/phuse/2015/CD/CD05\\_ppt.pdf](http://www.lexjansen.com/phuse/2015/CD/CD05_ppt.pdf)

PhUSE Wiki [http://www.phusewiki.org/wiki/index.php?title=PhUSE\\_Wiki](http://www.phusewiki.org/wiki/index.php?title=PhUSE_Wiki) (search for PRM)

Pines, Joshua "Protocol Representation Model in the Real World"

[http://www.lexjansen.com/phuse/2012/CD/CD06\\_ppt.pdf](http://www.lexjansen.com/phuse/2012/CD/CD06_ppt.pdf)

Vadakin, Andrea and Brooke Hinkson "Organizing and Accelerating the Clinical Research Process from the Beginning: The CDISC Protocol Representation Model and Toolkit"

[http://www.cdisc.org/system/files/all/article/application/pdf/protocol\\_representation\\_model\\_and\\_toolkit\\_vadakin\\_hinkson.pdf](http://www.cdisc.org/system/files/all/article/application/pdf/protocol_representation_model_and_toolkit_vadakin_hinkson.pdf)

## CONTACT INFORMATION

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

Jeffrey Abolafia [Jeff\\_Abolafia@rhoworld.com](mailto:Jeff_Abolafia@rhoworld.com)

Frank Dilorio [Frank@CodeCraftersInc.com](mailto:Frank@CodeCraftersInc.com)