Introduction to the Clinical Development Design (CDD) Framework

The PhUSE Semantic Technology Clinical Program Design in RDF Working Group

Table of Contents

Disclaimer

The opinions expressed in this document are those of the authors and do not necessarily represent the opinions of PhUSE, members' respective companies or organizations, or regulatory authorities. The content in this document should not be interpreted as a data standard and/or information required by regulatory authorities.

1 Overview

The Clinical Development Design (CDD) Framework encompasses design decisions and their attributes in clinical development programs and trials. These decisions rely on consolidating, analyzing, weighing, and prioritizing a multitude of information from guidance, past experience, expert advice, and real world evidence. This wealth of information is typically captured in an unstructured format, in the form of clinical development plans, protocols, documents, and presentations. The CDD Framework aims to provide structure and rigor to the design of clinical development programs consisting of clinical trials and clinical research programs. A cross-industry initiative under the auspices of PhUSE, with participants from industry, FDA, and academia, is working on an information model to underpin clinical development program design.

Pharmaceutical development organizations are faced with increasing challenges that are associated with the hurdles of successfully completing the myriad steps from drug development to product launch. Thus, they have started to explore ways to improve clinical development design decision-making, including the planning for, implementing, and analysis of clinical trials. Noting that automation aiding the execution, management, and reporting of clinical trials can only contribute to efficiency if the right trials are run, these organizations have turned their attention to ways of providing guidance, information enabling decision-making, retention of project knowledge and strengthening institution memory.

Currently the development and maintenance of clinical development design plans rely heavily on historical knowledge. This reliance is due to a lack of tools available to help create and recall the relationships between decisions and the information used to make the decisions. The further the decision is in the past, the more degradation occurs in historical knowledge, either due to key individuals leaving the organization or individuals forgetting the details of the decision process.

Technologies that easily link design components and decisions to the information/data are needed in clinical development to overcome the unavoidable loss of historical knowledge. Even with existing technologies that allow for creating relationships between design components and capturing “decision points”, individuals either only capture the final decision or capture the rationale and/or supporting data in text-based minutes of meetings.

2 Purpose

The workgroup members were charged with “Representing Clinical Program Design in RDF” or the development of Clinical Program Design tools. This document traces the evolution in our thinking on this project that required us to step back and understand the processes involved in Clinical Development Design (CDD) Framework, defines what we are calling the CDD Framework, and provides next steps in actually building tools and an Information Model to help

us.

The purpose of the *Introduction to the Clinical Development Design (CDD) Framework* and the activities of this working group are to focus the vision around the novel concept of structuring design thinking, rationale, and information underpinning clinical research. The paper itself is a vehicle that allows the concept to be fine-tuned through the synergies of cross-organizational thinking, and is aimed to further entice participation in this endeavor.

3 Scope

Information sharing would be facilitated by a standardized information model. Most of the current information modeling today focuses on individual clinical trials, and the representation of clinical trial data. Although work is ongoing to expand data standards to cover the protocol, these are insufficient to capture the objectives, rationale, and design thinking behind clinical programs.

The Clinical Development Design (CDD) Framework aims to provide structure and rigor to the design of clinical research, more specifically clinical development programs consisting of clinical trials and the work before clinical trials that are needed in implementation of the clinical research project. The CDD Framework explores the gaps and challenges in design practices supported by templates, documents, and unstructured tools. It sets out to first capture the language, terms, and vocabulary that we use to describe clinical development design. Even though the vocabulary may be quite similar, we try to differentiate the development design process from strictly protocol writing in terms of the decisions behind the choices made for the program. This vocabulary can then be used to build a simple Information Model. Although we eventually would like all of our work to be expressed in semantic technology and RDF cubes, our present scope is to identify basic terms and build conceptual maps describing the CDD process.

4 Definitions

4.1 Design and Decision-Making Terms

Term	Description
Information Model	An information model in software engineering is a representation of concepts and the relationships, constraints, rules, and operations to specify data semantics for a chosen domain of discourse. Typically it specifies relations between kinds of things, but may also include relations with individual things. It can provide sharable, stable, and organized structure of information requirements or knowledge for the domain

Term	Description
	context.[1]
Smart Program Design	Detailed process to improve the design and effectiveness of clinical development programs. The process requires data, historical information, and tools for improved information sharing. Currently, there are no agreed models across the pharmaceutical industry addressing the design of clinical programs. This program demonstrated commonality between companies and a joint desire to take the first step toward an industry-wide program design model in order to facilitate collaborations within and between companies as well as with vendors and authorities.[2]
Design Decisions	Kahneman suggests that at the design, production, and final inspection stages, we look for ways to improve our design decisions.[3]
Structured Design Thinking	1) Clarify and implement design approach and philosophy with the ultimate goal of gaining new insight for science, patients, payers and regulators. 2) Capture, bridge, and synthesize information from the multitude of inputs that underpin design decisions and future needs. 3) Assess and iteratively improve the practice of design.

4.2 CDD-Specific Terms

Term	Description
CDD Framework	Roadmap for developing a medical product, be it drug, device, or diagnostic
Stakeholders	Scientific/Medical, Operations, and Regulatory
Step 1 – Set Goals	Describe the product and target population - Endpoints connected to objectives and measures
Step 2 – Characterize Intervention	Determine nature of the intervention, dosing and administration: route, frequency, duration, and other parameters related to exposure. Focus on defining outcome measures - assessments linked by time and connected to the

Term	Description
	measures
Step 3 – Deliver a product	Demonstrate an acceptable benefit versus risk profile - attributes of the program design, population, trial design, comparator, drug and geographic location that support the collected measure. These define the limitations of the delivered product, e.g., what the data can support usage in.
Enabling Information	First, identify, collect, analyze, and archive the data critical to decision-making. Provide means to retrieve the data after archiving.
Enabling Data	• Regulatory information • Research and development/ decision-making • Clinical study data
Product Development Considerations	• To ensure careful and reasonable progress during development • To determine if the benefits and risks are acceptable for granting marketing authorization • To monitor safety during product investigation and following marketing authorization
Risk Management	The EMA definition of risk management as a systematic process for the assessment, control communication and review of risks associated with the planning and conduct of clinical trials and clinical development programs.[4] The CDD intends to expand this definition to include the entire life cycle of the product with special emphasis on the postmarketing phase.

4.3 Sources of Data to Construct our Database

Term	Description
Target Product Profile (TPP)[5]	A TPP is a format for a summary of a drug development program described in terms of labeling concepts which may be very helpful in its program design.
Quality by Design (QbD) Toolkit[6]	The Clinical Trials Transformation Initiative (CTTI) provides recommendations for improving clinical trials through a set of documents, templates, guidelines, and videos that help put

	QbD into practice.
FDA Checklists	One example is the "Selected Requirements of Prescribing Information," a checklist for prescribing information based on labeling regulations.[7]
ClinicalTrials.gov[8]	ClinicalTrials.gov is a registry and results database of clinical studies of human participants conducted around the world. The metadata behind the registry detail important study information.
Regulation2RDF (Reg2RDF)	Leverage the outcomes of the development work from the PhUSE Semantic Technology Reg2RDF workgroup.

4.4 Information Modeling Tools

Term	Description
Cmap[9]	The concept map or Cmap uses an open source software toolkit developed by the Institute for Human & Machine Cognition (IHMC) for the creation and manipulation of concept maps. It helps for organizing and representing our knowledge about a particular issue or system. In Cmap Tools, concepts (boxes or nouns) are linked together by propositions (lines or verbs) to form a network that visually demonstrates connections between issue components.
Neo4j[10]	Neo4j is a highly scalable native graphing tool, with a database component. It uses data relationships to help to build intelligent applications to meet today's wide ranging data challenges.
Semantic Technology[11]	• Natural Language Processing (NLP) • Artificial Intelligence (AI) and expert system/Knowledge-based systems • Data mining • Classification • Semantic search

4.5 Semantic Technology Terms

Term[11]	Description
Resource Description Framework (RDF)	A flexible data model
RDFS and OWL	Schema and ontology languages for describing concepts and relationships
SPARQL	A query language
RIF	A rules language

5 Problem Statement

The CDD Framework can provide a defined structure, capture links and data, and provide a cross-industry information model and terminology. Although these are disparate needs, they do fall under the umbrella of what a CDD Framework can do for clinical research projects.

While many design decisions may be supported by data, the link between those data and the decision is seldom, if ever, captured and stored. The link between design decisions and the TPP attributes is often just as ambiguous. How can reproducible design decisions be made?

The act of meticulously capturing design assumptions, constraints and supportive data throughout the product development lifecycle may seem to be an arduous overhead for the current project, given that the downstream cost for future projects can easily be discounted.

There is a clear need to maximise the efficiency and return on investment of the clinical development design process and to maintain the robustness, relevancy and currency of the metadata of CDDs in support of efficiency and maintenance of information capital. A common cross-industry information model and terminology for appropriately capturing the process of developing a CDD is needed. The CDD needs to accurately tell the story of how plans, data and information evolve from before entry into first in human dosing, followed by the clinic trial and entry into the market, as well postmarketing surveillance and studies.

6 Background

Due to the increased expenditure and difficulties associated with successfully developing and launching innovative new medicines, much effort has been spent on root cause analysis. Such analysis invariably reveals that clinical development has become more complex and fraught with significant risk on the return of R&D investment.[12] Furthermore, efforts to streamline

development processes, with automation and digitization of clinical trial data capture and reporting activities have failed to dent the overall success rate and cost of clinical research. The industry started to recognize the need to improve the design of clinical research.

Four key challenges have been identified in achieving consistent high-quality development program designs.[2] They are:
1. Design information is captured ad hoc.
2. There is an inability to learn from past programs, both within organizations and externally.
3. Current industry information standards do not cover program level or the rationale behind designs.
4. There have been limited opportunities for progress in improving clinical development programs due to the lack of new and collaborative tools.

7 Considerations

As can be seen from our original working group name: "Representing the Clinical Design Program" in RDF, we thought that we would link product design concepts and produce RDF cubes in a short period of time. Unfortunately, we did not even have a shared vocabulary or an ontology. Now we realize that we must consider a step-by-step process. Writing the manuscript, *The Clinical Development Design Framework,* we looked at the entire roadmap or framework for clinical development design and determined that the first step is to develop an ontology or set of terms from currently available sources, and start with the terms in Data Sources. Once we have a set of terms, we will design Cmaps or conceptual maps with our materials. We intend to have links within the Cmaps and use them as the basis for building our RDF Information Model.

As noted, information sharing is difficult. In terms of incorporation of legacy information, pre-existing information maintained by the organization must be assimilated, saved, and associated with newly collected information. The goal is to ensure the association of all information to justify prior and ongoing programs, and provide justification for new programs. Next we need to consider the ease of adoption of the tools and addressing what challenges are in place to make the transition, what new software is needed, and what software customization is required to fulfill the organizations objectives. And finally, we must think about the end user and the ease at which staff can transition to a new system, keeping the new technology as intuitive and simple as feasible.

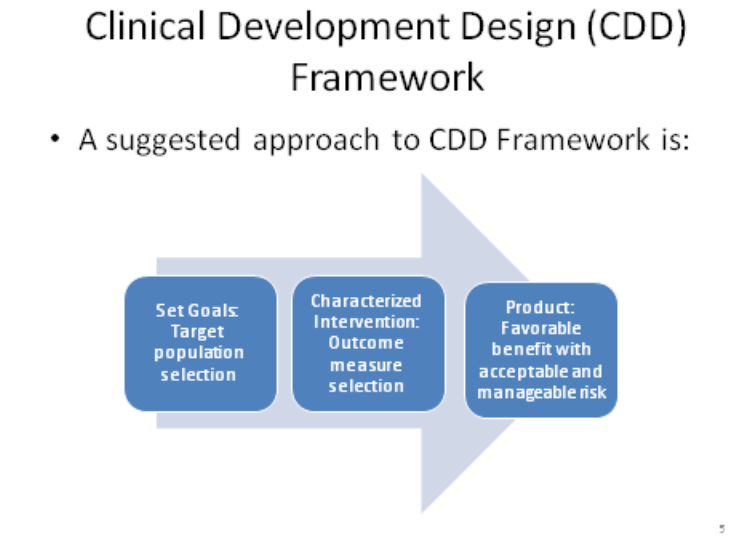
Before we can actually address the incorporation of the legacy information, we need a common vocabulary. What is included in Clinical Development Design? Who makes the decisions? What are the time parameters? Is our vocabulary applicable in all circumstances?

8 References

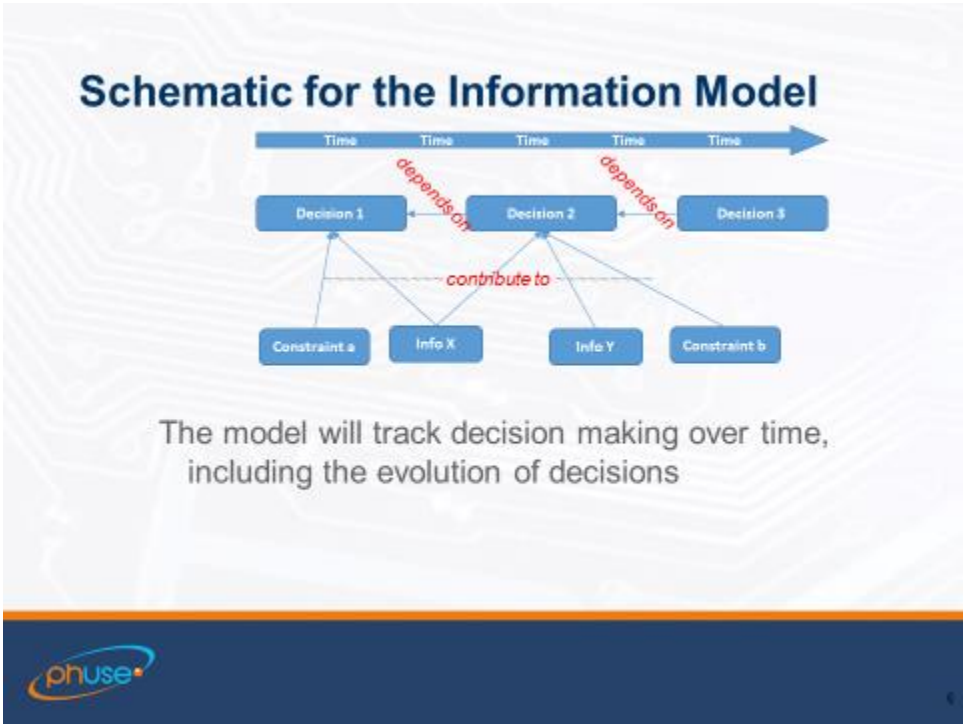
1. Lee Y. Information Modeling: From Design to Implementation. Gaithersburg, MD: NIST; 1999.
2. Vasko L, Sundgren M, Bachmann P, et al. Smart Program Design Through a Common Information Model. Therapeutic Innovation & Regulatory Science 2014.
3. Kahneman D. Thinking, Fast and Slow: Penguin Books Limited; 2011.
4. Proeve J. Clinical Study Risk Management: What Does it Really Mean & How do You Do It? DIA Annual Meeting; 2014; San Diego, CA.
5. FDA. Guidance for Industry and Review Staff: Target Product Profile — A Strategic Development Process Tool. Washington, DC: HHS FDA; 2007.
6. QbD (Quality By Design) Toolkit. 2016. (Accessed March 30, 2016, at http://www.ctti-clinicaltrials.org/toolkit/qbd.)
7. FDA. Selected Requirements of Prescribing Information, 21CFR201.56. 2009.
8. ClinicalTrials.gov. 2016. (Accessed July 13, 2016, at https://clinicaltrials.gov/.)
9. Cmap Concept Maps. 2014. (Accessed June 18, 2016, 2016, at http://cmap.ihmc.us/.)
10. Neo4j. 2016. (Accessed 6/23/2016, at https://neo4j.com/.)
11. Semantic Web vs. Semantic Technologies. 2016. at https://supportcenter.cambridgesemantics.com/semantic-university/semantic-web-vs-semantic-technologies.)
12. Reuters T. CMR International Pharmaceutical R&D Factbook. 2012.

9 Appendices

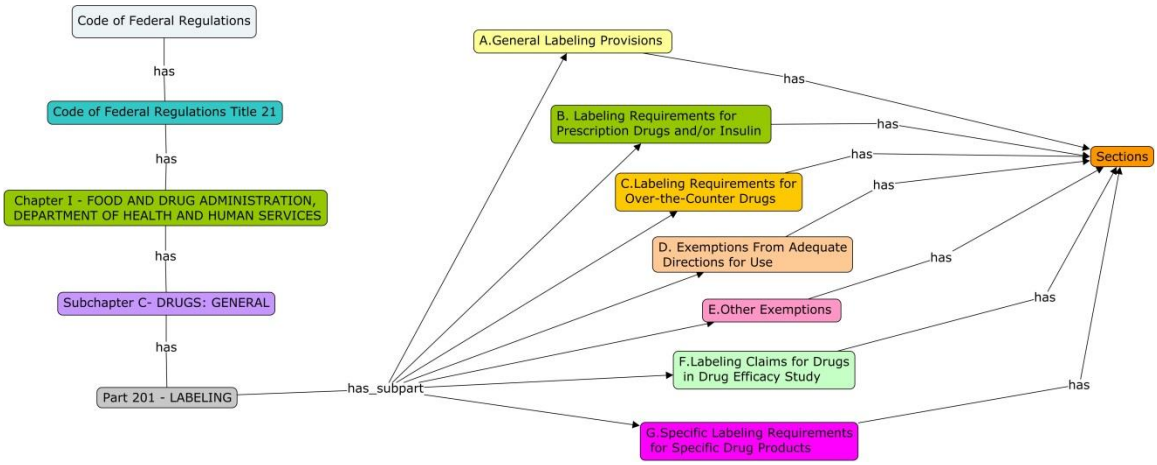
9.1 Appendix 1. Clinical Development Design Framework



9.2 Appendix 2. Ideas for Schematic



9.3 Appendix 3. Cmap of 21 CFR Part 201



9.4 Appendix 4. Neo4j – Visual Interactive Model