Appendix A: User Process Supported by Study Builder and Data Flow Schematics

Future State Narrative

Purpose

This appendix describes a technology-enabled business process for defining a clinical study design supported by a Study Builder platform. This narrative lays out the activities and functionality the tool should support. The final section of this appendix consists of schematics which explain the data flows implied by the business process.

The Study Builder will constitute the central component of a broader application ecosystem. Study Builder will articulate the study design by accessing centrally stored definitions from a Study Design Repository; it will enable the automated use/reuse of design elements in downstream applications by leveraging clinical data standards and study design metadata. The use of industry standards (e.g., CDISC, HL7), or company-specific standards and metadata, will enable structured protocol elements to be “prepared” for presentation to Application Programming Interfaces (APIs) that will in turn automate configuration of those downstream systems. The system configuration APIs themselves are out of scope for this document.

A Study Builder would be expected to support these two main capabilities:

1. Import, storage and accessibility of study design templates and associated standards.
2. Sharing or presentation of information that would support configuration of downstream systems.

Initially, the platform should prioritize information supporting configuration of electronic data capture (EDC) systems. Subsequently, the platform also should facilitate configuration of IRT, CTMS, eTMF and Portfolio Project Management systems. Eventually, other downstream applications could include clinical study registries, e-Consent and e-Source applications. Architecture of the tool should also support data inputs from other “upstream” sources of information such as site performance benchmarking data, study cost estimators, and other data sources that support study designers in determining patient populations and optimizing trial design.

The tool’s user interface should be based on structured prompts (e.g., selection menus, drop down lists, interrogative prompts, or a “wizard”) that will guide users through the workflow and design process.

In addition, the tool should support multiple output formats, including readable document formats (to populate the Protocol document itself) and digital formats according to the technical requirements of downstream systems. Each of those outputs should allow for the addition of supplementary data or specifications, or formatting for documentation. Any output data that requires correction (rather than supplementary information) should be highlighted back to the user, to prompt investigation of the associated standard, or to ensure links are correct.

A detailed listing of Key Functionality is provided (See Appendix C).

Supported Roles

The primary roles involved in the study design and build process include clinical science, biostatistics and medical writing; other functions such as data managers or system programmers may also have input to specific study design aspects. The narrative is written from a process or activity perspective to focus on the functionality the tool will support, rather than the specific job titles responsible for individual tasks. Because the process involves several roles, the tool should support collaboration among multiple users and should provide permission and access controls to avoid unintended changes.

Study Build Workflow

The workflow described below outlines how we envision different portions of the study being completed as “modules” that represent different aspects of the study design process. Information will be defined iteratively, likely by different authors, until all design elements have been defined.

The anticipated study design workflow is:

- Study Level Content
- Fixed/Common Study Design Elements
- Protocol Level Content
- Randomization Schema
- Visit Matrix
- Data Collection Elements

Start of Activity in Study Builder

Each company has a clear trigger that activates study design activities. In most cases, this is related to executive approval or budget approval. At this stage, high-level information (compound, therapeutic area, indication, protocol ID, phase and study type) may already be available in company systems (e.g., CTMS or Master Data).

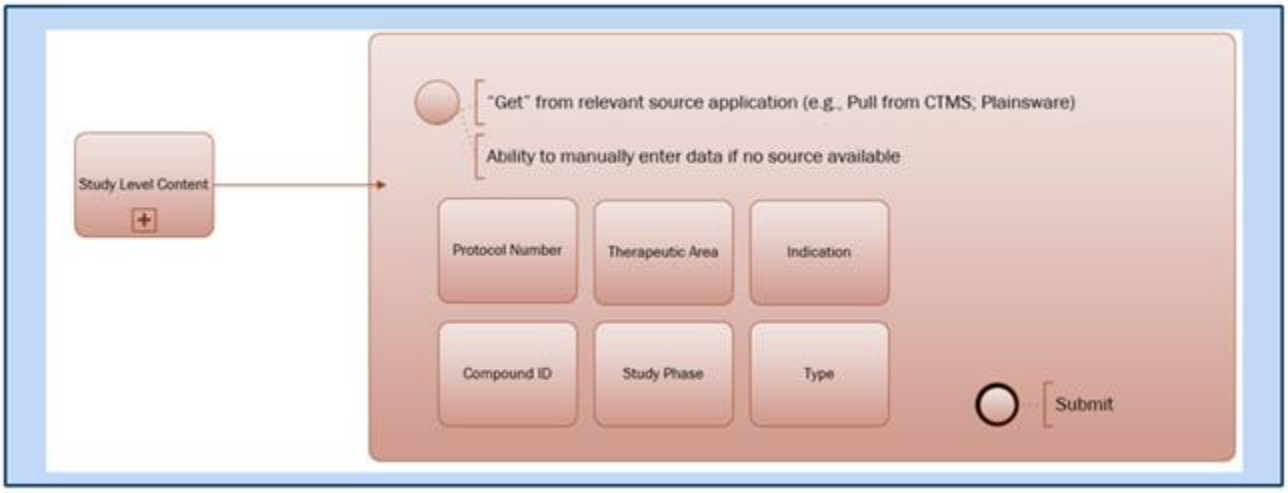
Study Level Content

Once study design activities are triggered, study build starts by entering protocol ID (study unique identifier), primary study treatment (compound), therapeutic area (TA), indication, study phase and study type (i.e., interventional, randomized, blinding, etc.). This information is most often available in source systems within the study sponsor (e.g., CTMS, Portfolio Project Management systems, or Clinical Development Plans) and the builder should be able to automatically upload the elements to be surfaced to the user upon prompting. If these data elements are not available in source systems, the user should be able to enter them into the Study Builder directly (See Figure 6).

Specifying TA, indication and compound would ensure subsequent steps are linked to the appropriate libraries of objectives, endpoints, assessments and forms. Subsequent choices will be directed toward the appropriate sets of standards by dynamic prompting to the appropriate sections of those libraries. Similarly, other information such as study phase, study type, study participant population, dose or study geography may further refine the appropriate libraries to be presented as study build progresses. For example, if study attributes indicate a phase 3, placebo-blinded oncology study of a particular compound versus standard of care in breast cancer, the library elements of objectives, endpoints and associated EDC and data collection elements for oncology/ breast cancer would automatically become prioritized options to the user.

For instances where a required standard or template does not exist (e.g., a new indication), the Study Builder would support creation of a new term and a workflow and approval process for that new term to be considered and incorporated as an additional standard in the Study Design Repository.

Figure 5: Study Level Specification



Fixed Study Concepts and Study Operational Requirements

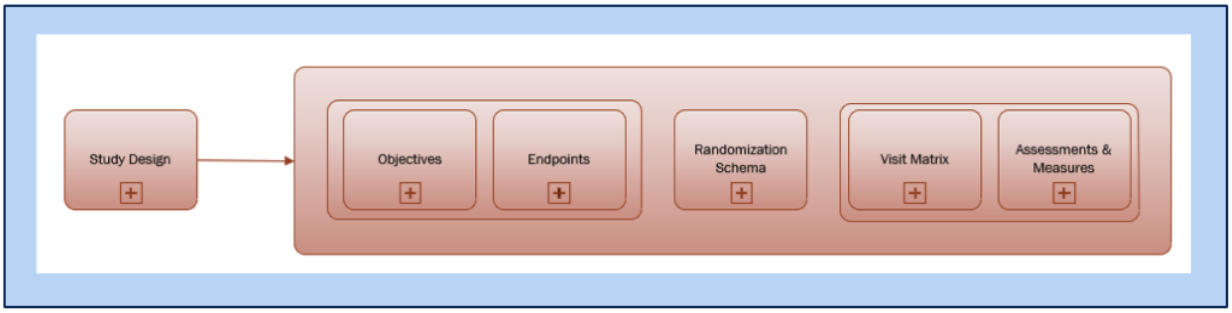
All studies share specific elements that are dependent on the TA or indication (objectives, endpoints, eligibility criteria and assessments) and many that are not dependent on the TA. The latter include baseline demographics, medical history, treatment preparation and administration, concomitant medications and adverse events (AE). The tool would ensure that all components are included in each complete study design. These non-TA dependent structural elements tend to be consistent across all studies, though specific versions of forms may depend on the therapeutic area (e.g., oncology-specific AE forms). Information regarding study comparator arms, randomization schema, study blinding, and treatment administration would also be entered in the tool.

Protocol Level Content

The protocol title, protocol short title and protocol acronym would be confirmed (if uploaded from a source system) or added. Each of these would be a redundant unique identifier linked to the protocol ID. Based on the TA, indication, compound, phase and other attributes, the tool will direct the user to choose from the most appropriate reference libraries of pre-defined objectives, endpoints and assessments.

Study Objectives would be selected/assigned from the relevant participant-, therapeutic area- or compound-specific content library, and categorized as primary, secondary, tertiary or exploratory objectives. Each objective would be linked to a TA-specific set of study endpoints which in turn is associated with biomedical concepts (clinical procedures, instrument or laboratory measurements) and definitions. Note that, in cases where the study being designed is identical or similar to another study (e.g., a prior study or a “sister study”) the tool should be able to access and pre-populate that information, with confirmation from the user after any required modification or adjustment.

Figure 6: Protocol Level Specification



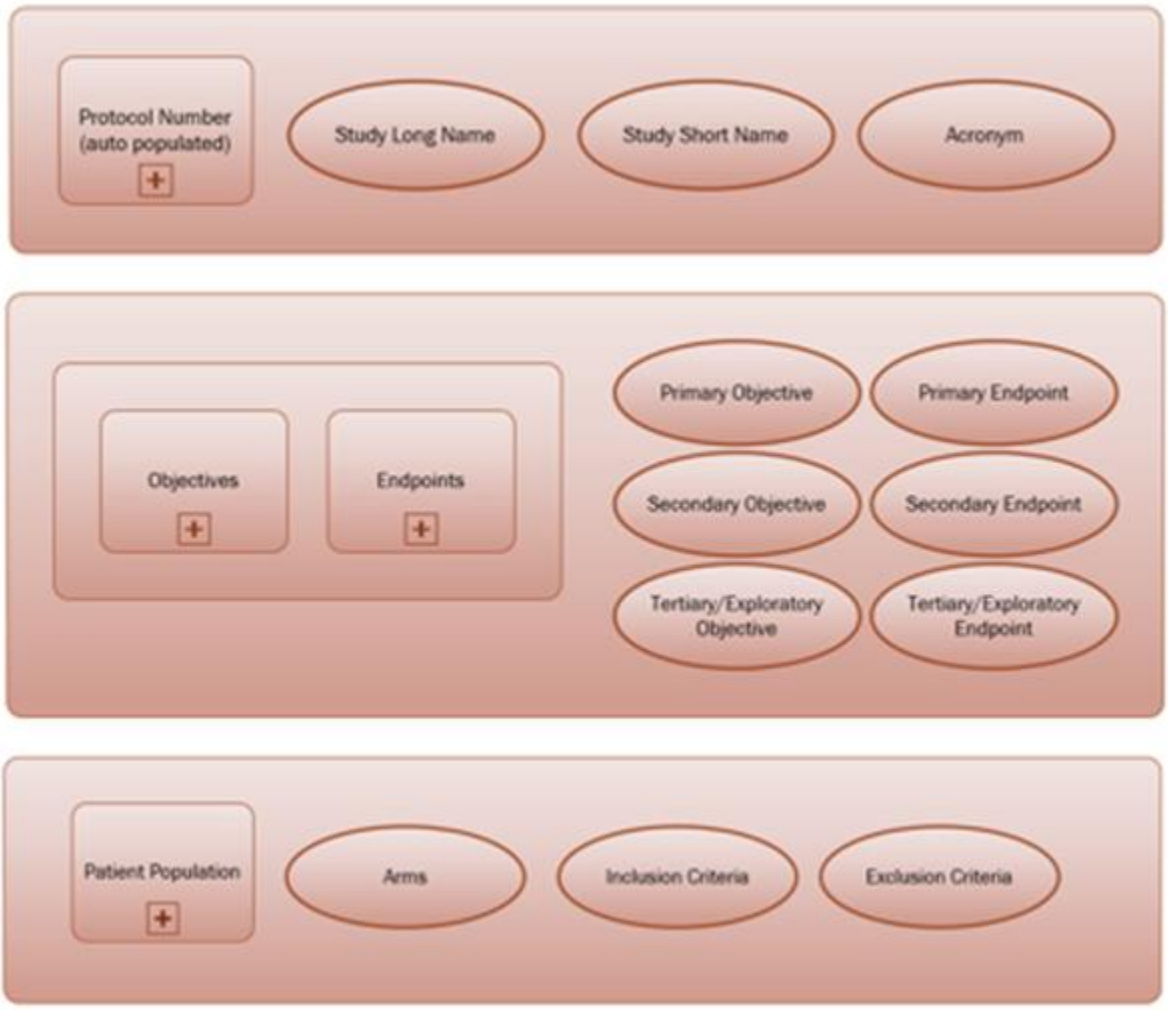
For example, using the TA Library for Asthma, a study in severe asthma could have as its Primary Objective "To evaluate the effect of drug x in participants with severe asthma." The primary endpoints linked to this objective are limited to "absolute change in percent of predicted FEV1 from baseline to [Week X]" OR "increase [magnitude of change] in FEV1 from baseline to [Week X]." This also implies that the FEV1 biomedical concept will require spirometry assessments to be scheduled at baseline (CDM: primary timepoint) and week X visits (CDM: secondary timepoint), and that FEV1 measurements will need to be captured in the study database, either by EDC or via data transfer. Further, options for Secondary Objectives include FVC or FEV1/FVC ratio (spirometry), reduction in symptoms (questionnaire data) or fewer Clinical Exacerbations (medical history or diary data) or reduction in the use of rescue medication (diary, dosing device or medication count data). As each objective is chosen, the appropriate choice of linked assessments and measures would also be assembled in the tool using the latest available standards for that assessment.

Randomization Schema and Study Design Concepts

The tool would prompt the user to specify the randomization schema and other elements of the protocol design. It would provide suggestions of study design elements based on the user's previous choices of indication, compound and other details.

The tool would support creation of new design elements as they are required and a workflow and approval process for that new term to be considered and incorporated as an additional standard in the Study Design Repository, for use in designing future studies. This intelligence is key to the vision for the tool.

Figure 7: Specification of Objectives and Endpoints



Visit Matrix Specification

The study visit matrix (or Schedule of Assessments, SoA) normally appears in table format, with rows representing the assessments to be completed and columns indicating the timing of the study visits or interventions. This representation serves as a central reference for the configuration of most study-specific tools and systems. The Study Builder should be capable to derive the study visit matrix by combining previously specified assessments with required timing, based on design templates associated with the therapeutic area, indication and study type or phase. Visit schedule should also be logically consistent with study objectives and primary efficacy or safety endpoints (i.e., if the primary endpoint is based on an efficacy assessment at Week X) and may be tied to pre-determined schedules (e.g., oncology studies

tied to treatment cycles). If standards or source data are not available in the Common Data Model repository, the builder will allow manual configuration.

Figure 8: Specification of Timing and Analytical details for Assessments

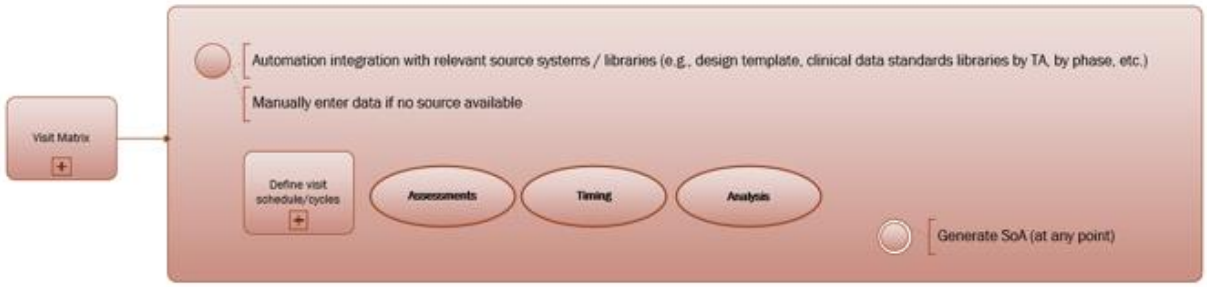
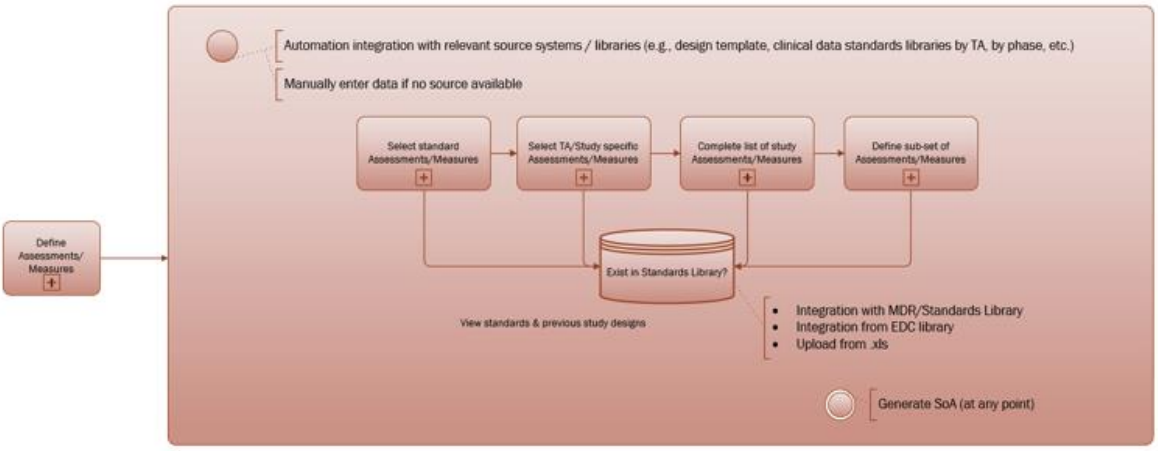


Figure 9: Selection (or Auto-Selection) of Standard Library Elements



To support configuration of downstream systems, specific details would be required for each visit, including, for example, limits for deviations from expected visit schedule ("visit windows" or expected study day for a visit, plus minimum and maximum allowable), and whether or not treatment is scheduled to be administered on that visit. In addition, planned analyses could also be associated with specific assessments.

Data Collection

Outputs of the Study Builder should link specified assessments to a current version of data collection standards. This would include standard EDC forms or file transfer specifications for non-CRF data.

The tool would allow specified assessments to be linked with relevant source systems or libraries (i.e., repository of design templates by TA or phase, data standards or CDM elements) that would link those assessments to standards and forms (such as EDC) and include any associated logic check specifications and validation scripts. These would also be linked to data

standards which will automatically annotate fields according to the applicable standard data model (e.g., CDISC ODM for systems configuration).

Version Control: Support for Protocol and Protocol Amendments

Once all details of study design are complete, and outputs that allow configuration of downstream systems have been created, the tool must allow outputs to be saved as an intact "version" of all structured elements for that study.

When the Study Protocol is subsequently amended, the tool should allow for specific sections or technical details of the design to be updated. Updates would be automatically linked to standards repositories. Outputs would be like the initial version, including both text and digital file outputs to update downstream documents or systems configuration, respectively. Once completed, the updated versions would be saved as a new, comprehensive version.

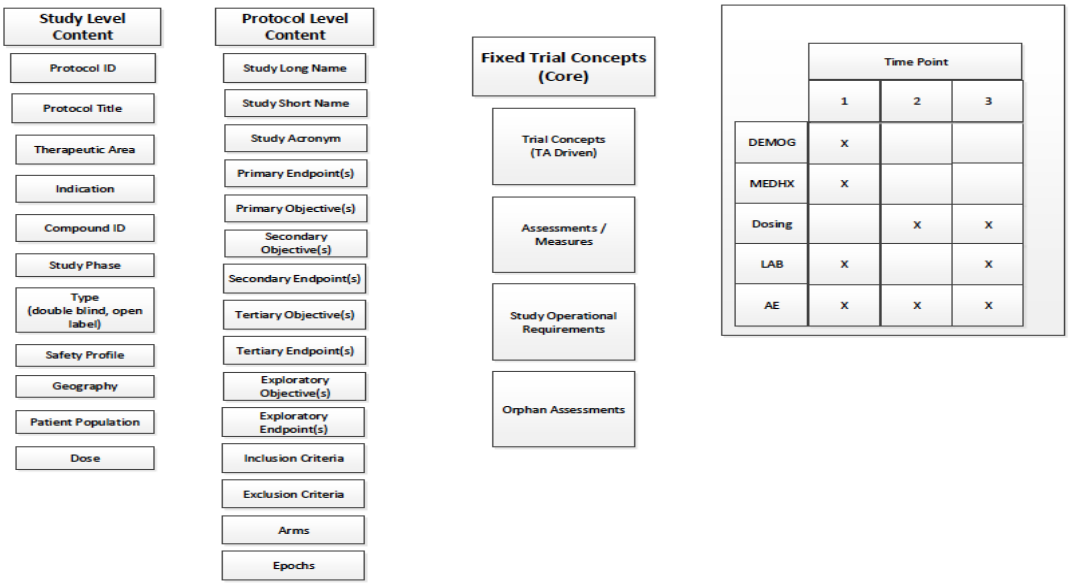
User Experience: "Real Time" Synchronization versus Discreet Updates to Downstream Outputs

As a protocol is created in the tool, it should be able to generate outputs in text and digital formats. This could be designed as discrete generation steps, where user command will cause the system to generate or render its outputs. Alternatively, the system could be configured to automatically create outputs continuously, in real time. The latter would be particularly beneficial for text outputs, as it would allow users early visualization of their work and facilitate early detection of errors or inconsistencies.

Data Flow Schematics

The following diagrams show the decision tree that would collect data essential for designing a study and feeding downstream systems. The sequence of figures demonstrates the logical flow described in the narrative, as the user moves from providing high level study descriptions (Protocol ID, TA, Compound) to more specific details about the study design (Item 1 below). The tool should prompt the study author to include “fixed” study concepts that are not study-specific but are included in every study (Item 2). The tool should also be able to derive the Schedule of Assessments from the structured information provided (Item 3). Items 4, 5 and 6 further describe how those data elements would flow through the process.

Diagram 1



1. **General Flow, Overall Data Elements:** The Study Level Content or Asset Level Content would represent the highest level of data required to plan a study or asset strategy. This information is typically stored at a strategic decision level – leading to a Go/No Go decision on an asset or a study. The data maintained here defines the essential study level data points that can be used to limit choices or libraries used to create study-level systems or plans. represents the details that would be required to begin drafting a specific protocol. By collecting this information as individual data points, the tool should not only be able to drive protocol authoring, but the data would be useable by other downstream systems, such as EDC.

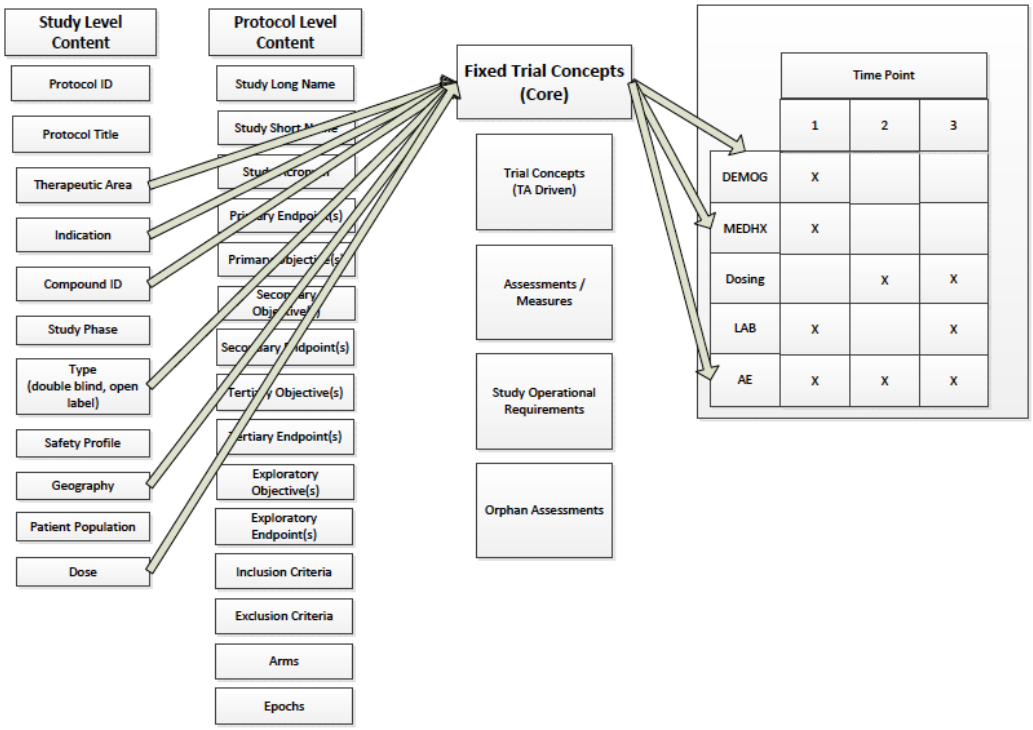
2. **General Flow Protocol, Fixed Study Concepts:** These are items that are minimally required to be included in every study. There may be some variations of particular items based on the study level data.
   a. **Study Concepts** include items that are minimally required to be included in every study of the design driven by the study level data selections. For example, “oncology” was selected as the Therapeutic Area, specific requirements for oncology

studies (e.g., specific adverse event or treatment administration forms) would be prioritized for selection.

b. **Assessments and Measures** are driven by the selected Objectives and Endpoints. Each endpoint and objective contain 3 pieces of data that can be used for decisions: Biomedical Concept being analyzed (e.g., lab value); Timing (e.g., baseline vs 12 weeks); Analysis Requirements (e.g., collected as a % value).
c. **Study Operational Requirements** are decisions that are required to support operational needs of a study. For example, an endpoint does not require the dosing information to be captured, but the analysis and study design will depend on this information.
d. **Orphan Assessments** are assessments that have been manually added to a study without being associated directly with a specific Objective or Endpoint, are not required for Study Operations and do not fit into the TA level requirements. These assessments would require some business approval to be added to the final study design.

3. **General Flow Protocol, Schedule of Assessments:** one of the key outputs from the information collected would be a digitized version of a SoA. Each point of required data would have been defined in the earlier steps.

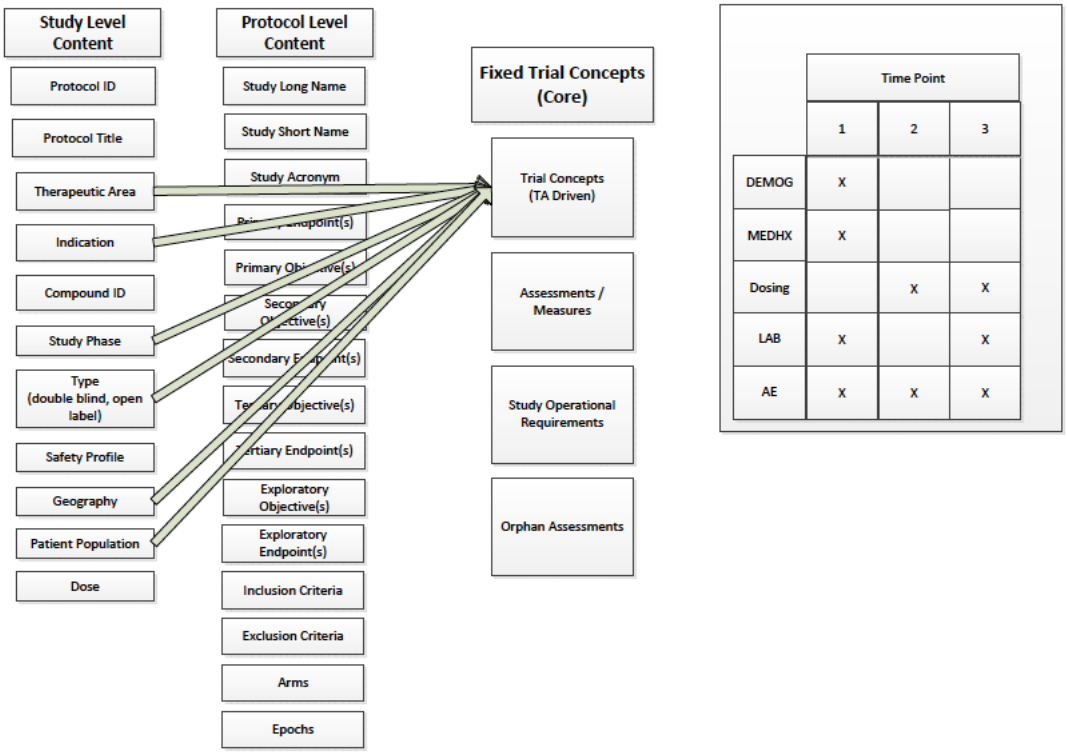
Diagram 2



4. **Fixed Study Concepts:** These represent “minimal required” information to run a study. They are common to all studies. For example, Patient Demographic information is

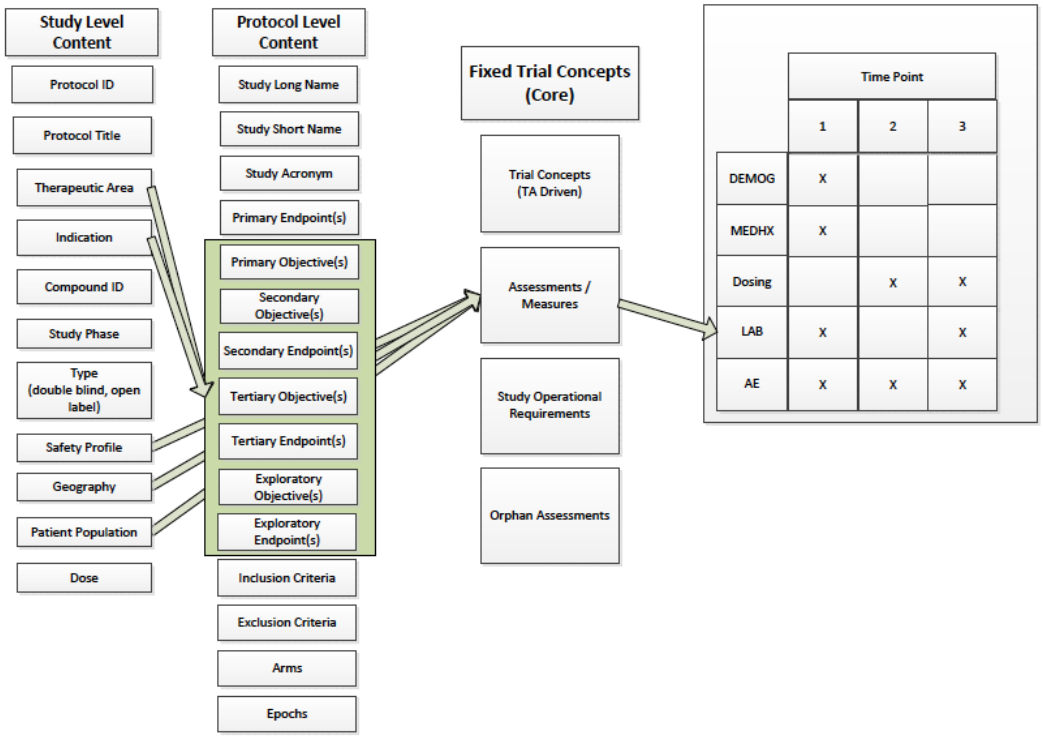
required in all studies with only minor variations, but the information collected in Study Level content can drive the available variants that are presented.

Diagram 3



5. **TA Driven Concepts:** Beginning at a “Study Level” or “Program Level” the tool would prompt high level questions that would limit the number of available libraries or data sets to select from in future steps. For example, using a specific TA will allow selections from any library (or standards repository) that are designated for that area. The Trial Concepts represent required material for the selected Study Level Content.

Diagram 4



6. **Assessments and Measures:** The Objectives and Endpoints would be selected from available libraries, limited by prior selections at the Study Level Content, and each endpoint will contain information that can be applied to three main areas: data collection, analysis requirements and timing in the study.