Journal of the Society for Clinical Data Management

Richardson A. A Case Study to Identify Required Schedule of Activity Characteristic Attributes. *Journal of the Society for Clinical Data Management*. 2024; 4(1): 7, pp. 1–12. DOI: https://doi.org/10.47912/jscdm.266

ORIGINAL RESEARCH

A Case Study to Identify Required Schedule of Activity Characteristic Attributes

Andrew Richardson

Introduction: The HL7 Fast Healthcare Interoperability Resources (FHIR) interoperability standard is now implemented widely to share healthcare records. The FHIR definitional resources are being investigated and developed to enable research study requirements to be used to support research data collection directly from the clinic to study sponsor. Schedule of Activities (SoA) as presented in clinical protocols require significant additional information before all requirements can be fully operationally implemented. This work has investigated the key characteristics required for SoA specification.
Objective: The objective was to identify a minimal set of SoA attributes that enable a machine-readable definition to communicate a study's scheduling requirements automatically and accurately. The FHIR "PlanDefinition" definitional resources as defined by the HL7 Vulcan SoA Implementation Guide were used as the target output for automating SoA generation and confirming specification accuracy and completeness.
Methods: Using scheduling and workflow principles, a minimal set of attributes to model SoA requirements was identified, tested, and developed using graph methods. Standard and more complex SoA types were used to identify and test those attributes appearing in all or most SoAs. Graphs with the identified attributes were developed, reviewed, and iteratively improved for their ability to (a) describe accurately the SoA, (b) enable manipulation using standard graph methods, and (c) be generated as FHIR resources compliant with the recently published HL7 Vulcan SoA Project Implementation Guide.
Results: A minimum viable set of SoA characteristics has been identified that can describe SoA requirements across a range of different study designs. These include simple linear schedules, cycles (eg. as in oncology trials), and event driven (eg, as in vaccine trials). In the simple linear and some of the more complex designs these could be generated automatically as SoA FHIR Implementation Guide (SoA IG) compliant resources. Some SoA requirements, such as conditional paths, whilst easily represented in the graph model, were not so easily represented as SoA IG resources, highlighting some current limitations with this interoperability standard.
Conclusion: A minimum viable set of SoA characteristics able to describe common study timing requirements for defining, creating, and manipulating SoA requirements was developed using graph methodologies. The primary focus was on generating study-specific FHIR resources compliant with the HL7 Vulcan SoA IG for communicating study SoA requirements to investigatory sites. Although this work was aimed at defining SoAs in FHIR format, the same attributes should be present and identifiable in other SoA models, such as the DDF USDM. The findings here, therefore, may have a broader applicability for confirming machine-readable SoA requirements.

Keywords: Clinical Trial; Schedule of Activities; eSource; Direct data capture; Fast Healthcare Interoperability Resources (FHIR); Operational implementation; Machine Readable

Introduction

The protocol schedule of activities (SoA), which describes the progress of a research participant through a study, is at the heart of each clinical trial. While they are most commonly presented in tabular form, they can take on various forms. The SoA usually only serves as a good initial specification for subsequent operational implementation, and will require considerable review, interpretation, and confirmation before all study details are understood. The interest in automating key operational outcomes driven by the SoA (eg, configuring electronic data collection (EDC) systems) and the interest in protocol content reuse to support digitisation and operational efficiencies has focused efforts on developing better methods to describe, and subsequently work with SoAs, particularly

Zenetar Ltd, Ramsbury House, Charnham Lane, Hungerford, Berkshire, RG17 0EY, UK
andy.richardson@zenetar.com

in machine-readable format (eg, the Common Protocol Template)[1,2]. These initiatives are largely focused on improving 'downstream' efficiencies for clinical study teams or sponsors. Direct data extraction from electronic health records (EHR)[3,4] and the use of real-world data (RWD)[5–7] to support clinical research is highlighting the limitations of the protocol SoA for specifying 'upstream' requirements.

Background

Turning study requirements as described in the protocol and the protocol SoA into operational tools and systems has remained largely unchanged for the last 30 years. Each operational group – clinical operations, data management, central laboratories, supplies and randomisation management, etc. – take the protocol SoA, interpret and expand on it as required by their function, and implement those parts for which they are responsible. In sponsor settings, standardisation and the tight integration of systems may assist with overall implementation efficiency (for example, the use of sponsor standards for configuring and integrating a Clinical Trial Management System (CTMS), an Interactive Voice Randomisation System, and the EDC, etc.). Even in this case, few studies proceed without additional study specific requirements needing to be put in place.[8]

In its tabular form, the SoA generally reflects the importance of each activity at defined timepoints for the experimental outcomes under investigation, as decided by the protocol authors. Often, between studies, the SoA may have highly variable information. In many cases it may serve as a summary of requirements that are described elsewhere; at the other extreme, some authors load the SoA with all the information they consider necessary to cover all protocol scheduling and activity expectations. **Figure 1** illustrates some of these points.

However, the SoA still remains the primary starting point for configuring most clinical applications such as EDCs, CTMSs, etc. Application implementations often adopt the easy-to-understand tabular SoA format (often with drill-down menus or similar tools) to assist with application navigation and/or data management tasks. This requires that a description of the SoA is held internally. Changes to the protocol, and specifically changes to the SoA, then require updates to the internal representations. These updates are often not automated and will require manual intervention (eg, using spreadsheets) before implementation.

The limitations of the tabular SoA format for automating operational requirements are further challenged when considering how to configure external applications, such as EHRs for direct (clinical trial) data capture, or for Real World Data (RWD) data collection, in which the research SoA is not a *de facto* functional requirement implemented by such systems. The increasing interest in adopting the HL7 FHIR resources as a standard approach for defining clinical research plans offers new ways to communicate trial requirements to participating organisations independent of the application functionality.[9–12]

The aim of this work was to systematically review the core objectives of the SoA and the use cases it supports with the goal of developing a generic model that can (a) describe any reasonable operational requirement specified by a protocol SoA; (b) incorporate these into a model with an appropriate level of granularity to be operationally useful; (c) be extended for additional SoA use cases, such as identifying user roles; and (d) be able to be used for the generation of application-specific specifications, if required. The operational focus for (d) was to understand the specific requirements needed to define, manage, and then represent SoAs using the HL7 FHIR interoperability standards.

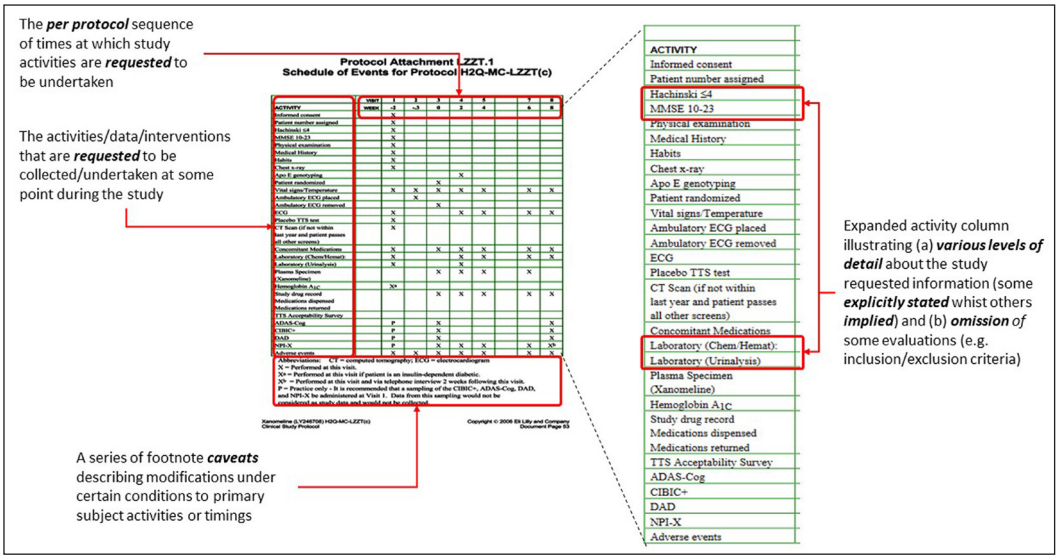

Figure 1: A study schedule of activities (SoA) annotated to highlight **(a)** the primary operational contribution of the principal components and **(b)** the variability in detail. This example summarises the activities for an ***enrolled*** subject only – hence the omission of any inclusion/exclusion criteria details – which are detailed in the text of the protocol, as are expanded details of each of the planned activities. Example is from the CDISC Pilot Study.

Methods

The development of the model used a three-step iterative process as described below:

1: Systematic review of the purpose and key information in SoAs

More than 40 clinical study protocol SoAs from pharmaceutical randomised clinical trials (RCTs), academic investigations, and registry surveys were systematically reviewed to extract the details of the study information that they were designed to communicate. Only studies in which the full protocol was available were used; these included studies in Alzheimer's, vaccine development, diabetes, and other therapeutic areas. An indicative set of the type and range of protocols reviewed is shown in **Table 1**.

SoA tables were selected as the starting point as they offer the best overview of a study's interventions and data collection requirements and are relatively easy to reflect programmatically. Protocol text with SoA information was not excluded from these reviews and was used to test the scope, design, omissions, or duplication of information initially identified from the tabular SoA presentations. Use cases/stories[13,14] that answered the generic question/test "*who* (...is using the SoA) to communicate *what* to *whom* and *why* (...ie, for what purpose)?" were developed. Variations on the primary theme were then tested by review to identify common features (eg, by identifying which roles use the same information for similar reasons).

The results of these reviews were used to

(a) identify the protocol components that contain scheduling information,
(b) recognise primary and secondary use cases and associated roles,
eg, "as a [data manager] I use the SoA to [develop data checks]", "as a [study participant] I [indirectly] use the SoA to [have my study appointments scheduled]"),
(c) identify the information required directly and indirectly (ie, not available or found elsewhere) in the SoA that would be required for operational implementation.
eg, 'Haematology' implies the measurement of a specific set of blood parameters, requires blood samples to be drawn, for those samples sent to a laboratory and analysed, and for results to be returned to the site/sponsor, and requires several skilled resources in several different study roles for completion.

2: Development of a SoA Characteristic Attributes Model

Using the results of the reviews in (1), a SoA graph model with general applicability to a wide range of SoA types was developed to (a) reflect protocol SoAs accurately, (b) edit and extend SoAs to incorporate other operational requirements and additional use cases, whilst (c) ensuring exports of the model in any format remained consistent.

Table 1: Representative sample of the type and range of study protocols reviewed. 40+ protocol SoAs were reviewed in detail and others used informally to confirm review outcomes.

Study Identifier	Study Title
CDISC Pilot Study	Safety and Efficacy of the Xanomeline Transdermal Therapeutic System (TTS) in Patients with Mild to Moderate Alzheimer's Disease (LZZT)
NCT04320615[a]	A Study to Evaluate the Safety and Efficacy of Tocilizumab in Patients With Severe COVID-19 Pneumonia
NCT04505722[a]	A Study of Ad26.COV2.S for the Prevention of SARS-CoV-2-Mediated COVID-19 in Adult Participants
NCT04193176[a]	Efficacy and Safety of Gefapixant (MK-7264) in Women With Chronic Cough and Stress Urinary Incontinence (MK-7264-042)
NCT04368728[a]	Study to Describe the Safety, Tolerability, Immunogenicity, and Efficacy of RNA Vaccine Candidates Against COVID-19 in Healthy Individuals
NCT03653546[a]	First Line Treatment in EGFR Mutation Positive Advanced NSCLC Patients With Central Nervous System (CNS) Metastases
NCT05502692[a]	CHARACTERISE – A Cross-sectional, Observational Study to Characterise the Transition to Dolutegravir-based Regimens in South Africa in Terms of the Emergence of Obesity, Viral Re-suppression and Integration Into Routine Programme Care (CHARACTERISE)
NCT06141343[a]	Project V – A Randomised Controlled Prospective Study of the Next-generation Probiotic, Veillonella Atypica FB0054, vs Placebo in Healthy Adults
NCT04470427[a]	A Study to Evaluate Efficacy, Safety, and Immunogenicity of mRNA-1273 Vaccine in Adults Aged 18 Years and Older to Prevent COVID-19
NCT02328768[a]	Compassionate Use of Omegaven® for the Treatment of Intestinal Failure Associated Liver Disease in Children
ISRCTN72331636[b]	The OPAL Study: Older People And n-3 Long-chain polyunsaturated fatty acids

[a] Accessed through *ClinicalTrials.gov*.
[b] https://doi.org/10.1186/ISRCTN72331636.

The review cycle details were used as a design starting point; thereafter each design was iteratively tested to ensure any model limitations were recognised and compensated for or re-designed as appropriate. This was particularly important with regards to both the level of granularity and the range of SoA information that could reasonably be incorporated into the design without compromising the overall design objectives.

3: SoA Representation as FHIR Resources

The HL7 Vulcan Schedule of Activities Implementation Guide[21] (and principally the FHIR PlanDefinition, ActivityDefinition and related definitional resources) were used to identify those FHIR resource elements which are necessary, mandated, optional and use standardised codes to represent/generate SoAs accurately in this format.

4: Model Testing and Proof of Concept (PoC)

The SoA model from (2) was implemented using graph database methods to develop and test how the SoA key characteristics could be accurately represented and manipulated.

At each iteration the review findings and the model were tested by developing proof-of-concept (PoC) examples using the Python[15] generalised programming language, the NetworkX graph and network libraries,[16] and the pandas data analysis library.[17] PoC FHIR Resource examples were generated using the Python fhir.resources library[18] and the HL7 FHIR Shorthand (FSH) utilities and methods.[19] The yED graph editor[20] was used to create the visual graph presentations.

The accuracy of the SoA FHIR Resources versus the graph model and SoA tables were quality controlled by visual comparison, and data requirements and coding logic were revised until no errors existed. PoC examples were developed (a) *de novo* to confirm basic model designs, (b) based on examples in part or whole from reviewed SoAs (eg, to test linked SoA tables), and (c) using the Clinical Data Interchange Standards Consortium (CDISC) Pilot Study protocol (Eli Lilly LZZT Alzheimer's Study) as an example of a complete study.

The resulting FHIR Resources generated from each PoC example were confirmed as valid FHIR Resources by loading them to publicly available FHIR endpoints, recovering the specifications using FHIR searches, and thereafter confirming that the full original specifications could be recovered without information loss.

Results

SoA Review – Schedule and Activity Identification

Figure 1 shows the four general components that were recognised in all the SoA types (as tables, diagrams or other designs) reviewed during this work. The four key components are:

- A ***per protocol*** sequence of times and timings at which requested activities are to be undertaken. This usually represents the schedule for an ***ideal participant who completes all study activities as planned.*** The use of multiple SoAs that detail particular sub-schedules and activities is also common.
- The range of ***activities*** that are requested to be undertaken at some point during the study, defined to various levels of detail.
- ***The matrix of 'X's*** that define at which scheduled timepoint a specific activity is to be undertaken. Sub- or super-scripted 'Xs' are regularly used to highlight variations on the theme.
- ***'Variations on the theme'***, which are recognised modifications to either the schedule or the activities under identified circumstances, and are usually presented as footnotes to the SoA table or as references to protocol text (however, depending upon the authors approach, may themselves be detailed as tables or other diagrams).

Table 2 shows examples of the common protocol SoA elements that define scheduling requirements, together with some usage examples. While the detail provided in the protocols is assumed to be ***sufficient for the purpose of the study***, large variations in the level of detail, and their presentation, exist. For example, some SoA tables were complete without further notes, while one protocol reviewed included two SoA tables with a combined total of 56 supporting footnotes.

Many differences in the approach to detailing important operational or regulatory requirements were found. For example, when, what or if to record 'unscheduled visits' was sometimes included in SoA tables, sometimes in text, or was sometimes absent. Similarly, identification of other ***contacts with study participants*** (not necessarily 'visits') using the general form ***'repeat [activity] each day for 7 days'*** (eg, follow-up telephone calls) was observed regularly, but was highly variable in how and where it was specified.

The use of ***several SoA tables*** is a regularly met feature, especially in circumstances in which the study involves complex scheduling, such as oncology (eg, to detail cycles of treatment) or where progress through the study is dependent upon different circumstances (eg, vaccines/immunisation, or assignment to study arms with different schedules).

SoA Review – The Meaning of 'X'

The use of 'X' – and what it is meant to communicate within SoAs – is highly variable.

The uses of X can range from the simple 'undertake the task/activity at this point', through instructions under certain conditions (eg, 'only IF patient is…') to links to other (sub) schedules. The level of 'activity' detail in an 'X' can also be very variable, and can range from explicit/clear descriptions, to implied activity or even not be stated. These methods assume the reader will inherently understand how to interpret the intent. **Table 3** shows examples of the meaning of 'X' from the CDISC Pilot Study SoA (Lilly Xanomeline Clinical Study Protocol – LZZT) in which the implied or stated meaning of 'X' ranges from a request to record an observation (eg, height), through requests for an expert determination of a change in subject circumstances (eg, review concomitant medications or adverse events), to instructions to follow a 'mini schedule' (eg, workflow), such as that required

Table 2: Protocol elements/components that regularly contain scheduling and activity details with typical examples. The key SoA review test applied in each case was "is the [element/component/section/text/etc.] [representing/specifying/detailing/describing...] a planned activity?". For this exercise, a "planned activity" was defined as an activity with a specified study timing/scheduling element together with one or more associated tasks.

Protocol Element	Type of SoA Content	Example(s)
SoA Table (Single)	Primary planned study activities	*Per Protocol* participant path through study
SoA Table (Multiple)	Conditional or special case planned study activities	Vaccine study disease evaluations Variable oncology treatment regimens Treatment arm variations Additional special evaluations
SoA Matrix Symbols	Study required scheduled activity Modifications to activity (supported by legends/footnotes) Activity information	"X" "Xf" "17.5" (mL blood draw)
SoA Footnotes	Alternative or conditional paths through study Repeated evaluations specified once Conditional activities Type of scheduled "visit" "Workflow" schedule	"Only if" conditions "Every week for 12 weeks" "Xf" Female only" "P" procedure training "Telephone call"
SoA References	References to activity details References to scheduling frequency	Procedure description(s) *"---- Continuous ----"*
Protocol Sections, Non SoA tables, Appendixes	Procedure timings Timing caveats Non Case Report Form (CRF) data collection instruments, methods	Sampling for pharmacokinetic analysis Measurement frequency (3×/day) Timings for allowed or disallowed interventions, medications, etc. ePRO, eCOA recording frequency.
Workflows, Standard Operating Procedures Clinical Procedure Descriptions Other documentation	Descriptions of the required timings and tasks for specified activities (not described in protocol)	Regulatory reporting requirements Device data logging frequency

Table 3: 'Meaning of X' examples illustrating typical protocol SoA table use of 'X' together with the required operational outcome. The examples are from the CDISC Pilot Study SoA table of **Figure 1**. The examples show that to define 'X' unambiguously in machine-readable formats, various degrees of interpretation and further information are needed in many cases.

Activity	'X'	The Meaning of 'X'	Notes about 'X'
Height	X	Measure the subject height and record in the electronic Case Report Form (eCRF)	Simple observation
Vital Signs	X	Measure the BP and pulse rate and record in the eCRF	Details elsewhere "see section * for more details..."
Laboratory (Haematology)	X	Take a blood sample for analysis and send to lab. Analyse the blood and report to site, study... Make the results available in the study database	Multiple actors involved (site, laboratory, data management). Required tests specified elsewhere in protocol
Concomitant Medication	X	Review the subject's medication status AND IF CHANGED record the revised medication details in the eCRF	Study and/or expert knowledge required
Adverse Events	X	Determine if the subject has experienced any events that require reporting to the sponsor AND IF YES, record the events in the eCRF	Study and/or expert knowledge required
[Various, with SoA table footnotes]	X[a], X[b], P...	IF condition/event applies to subject THEN modify the standard schedule or activity as per the protocol	Understanding of alternative study schedules/activities required (eg, for M/F)
Plasma Concentration	X	Take blood samples for analysis, as per the sampling schedule	Timing details elsewhere in protocol "See Appendix 2"
[Protocol Text]	n/a	BP to be measured 5 mins after supine in dominant arm...	Observation conditions detailed elsewhere in protocol

for obtaining samples for pharmacokinetic analysis. It is not uncommon to find 'X' replaced with constructs such as '←===→' or '=== *continuous* ===',[2] with its detailed scheduling/activity requirement being only being clear in context. For example, 'Continuous' ECG recording is not the same as 'Continuous' adverse event monitoring.

SoA Review – Who, What, for Whom, and Why?

Table 4 shows a range of identified use cases in which the SoA is key to communicating study requirements and illustrates the wide range of roles and reasons in which SoA information is required for accurate operational implementation. It also gives some insight into where and what additional operational information is required for successful implementation if using machine-readable/consumable methods.

SoA Characteristic Attributes – Object Model

Figure 2 shows a subset of the CDISC Pilot Study SoA visits represented as a directed graph (specifically a NetworkX DiGraph,[16] which store nodes and directed edges with optional data/attributes). This offered a reliable approach for defining and manipulating SoA details. SoA 'visits' and 'activities' were represented as nodes in the model, with the SoA sequence/order represented by the edges. It was helpful to categorise nodes as one of two basic types – 'Interactions' and 'Activities' – that were used to distinguish the fundamental SoA objects.

Table 4: Selected illustrative examples of "***Who*** is communicating ***What*** (using the schedule of activities) to ***Whom*** for *what* ***Purpose (Why)***" placed into four categories (Study Objectives and Regulatory Approval, Project Oversight and Operational Requirements, Operational Functional Requirement Planning, and Operational Functional Implementation). These illustrate the wide general areas that draw on the SoA for study details and highlight the hierarchical cascade that occurs from initial protocol drafting and approval to final operational implementation, which adds additional requirements. The table is not exhaustive; other use cases in which the SoA is a primary specification source are left to the reader to identify.

Who	Communicating What	To Whom	For (Purpose)
Study Objectives and Regulatory Approval			
Study Sponsor (Clinician, Statistician, Medical Writer)	Proposed study participant schedule of evaluations and measurements	Regulatory authorities, Ethics committees	Study review and evaluation, Study approval
	Permitted or required variations to the schedule or evaluations and measurements (eg, caveats, footnotes, etc.)		
Project Oversight and Operational Requirements			
Study Sponsor (Project Manager)	Study participant schedule of evaluations and measurements	Sponsor project teams, Site study teams, etc.	Evaluation and measurement timings, ie, 'Visits'
	Required study evaluations and measurements		Evaluations, Measurements, Interventions, Instructions, etc. ie, 'Activities'
	Per protocol participant planned schedule of evaluations & measurements		Specification of study primary path or path(s)'
	Recognised alternative participant schedule(s) of evaluations and measurements		'Unscheduled Visits', Conditional activities
	Period for recording relevant medical history, medication details, etc.		Feasibility, Inclusion/Exclusion Criteria
	Permitted schedule timing variations		'Visit Windows'
	'Visit' type		'Clinic Visit', 'Telephone call', 'Remote consultation', etc.

(Contd.)

Who	Communicating What	To Whom	For (Purpose)
Site Study Teams (Study Nurse)	Identification of required resources (type, qualifications) to complete activities	Clinical staff, Hospital departments, Clinical Laboratories	'Primary Investigator,' 'Sub Investigator; 'Nurse Practitioner,' 'Radiographer', Equipment, Procedures, Diagnostic Services, etc.
Operational Functional Requirement Planning			
Study Sponsor (Project Manager)	Required study evaluations and measurements	Sponsor project teams, Service provider project teams	Operational implementation
Sponsor project teams (Functional Leads)	Required study evaluations and measurements	Functional responsible roles (clinical operations, supplies management, clinical data management, statistics, regulatory affairs, etc.)	Study application configuration, development, build, testing, etc.
	Required study evaluations and measurements	Study Lead DM	Resource requirements
Operational Functional Implementation (Data Management)			
Study Lead DM	Required study evaluations and measurements	Project DM	Data checks and review development
	Required study evaluations and measurements	Coding Groups	Coding requirements
	Randomisation timing	IVRS Support	Application integration

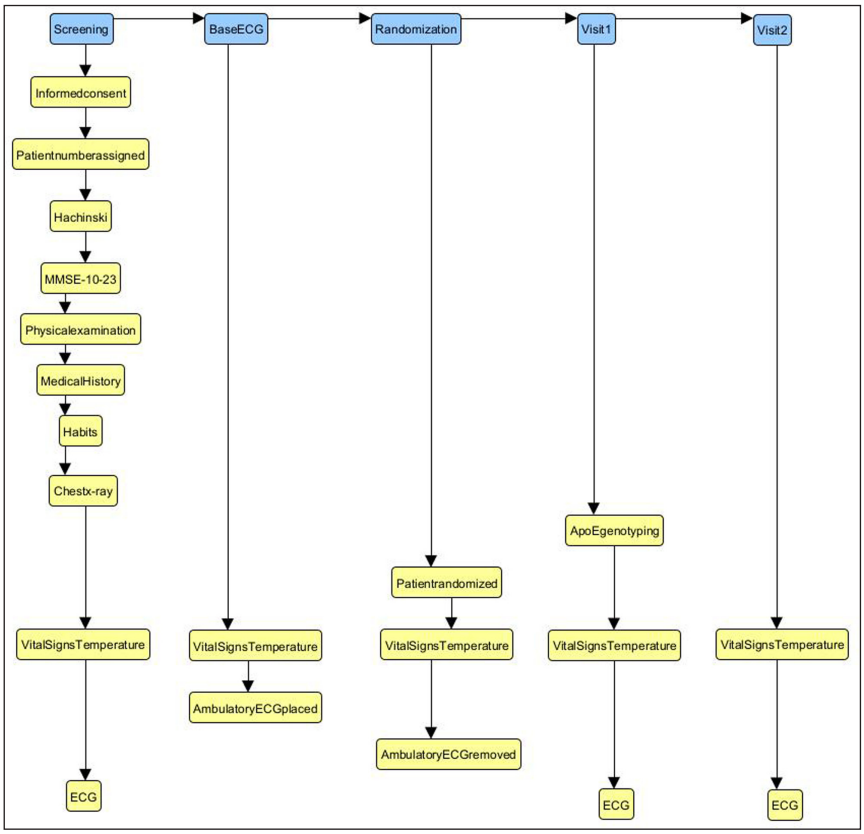

Figure 2: The first five 'visits' and thirteen 'activities' (top left corner) of **Figure 1** represented in graph form. The blue nodes in the diagram represent the '**interactions**' as described in the text, and the yellow nodes the '**activities**'. The relationships between the nodes shows the order as presented in the protocol SoA table. See text for further details.

SoA Model 'Interactions' (**Figure 2**, blue boxes) are defined as:

> *"a communication or involvement, either directly or indirectly, of the study SoA sponsor with study team members and/or research subjects or participants".*

SoA Model 'Activities' (**Figure 2**, yellow boxes) are defined as:

> "*a set of study tasks and/or requirements to be executed or satisfied contiguously".*

This level of abstraction was used to enable systematic testing of the PoC examples against the original SoA specifications (however defined).

Minimal Viable SoA Characteristic Attributes

Table 5 shows the final set of characteristic attributes necessary to represent and manipulate SoA requirements prior to re-representation as Vulcan SoA Implementation Guide (IG) compliant FHIR resources.

The study showed that the level of protocol detail considered sufficient to describe any specific schedule is (a) variable, (b) often located in different parts of the protocol, and (c) may not actually be stated (but implied through convention). For example, *interaction subtype* details (eg, visit type) might be found in a SoA table row, sometimes in footnotes, or only in protocol section text. The basic relationships between the three SoA Graph Model objects (Interaction Node, Activity Node, Interaction/Activity Edges) and an SoA table is shown graphically in **Figure 4**.

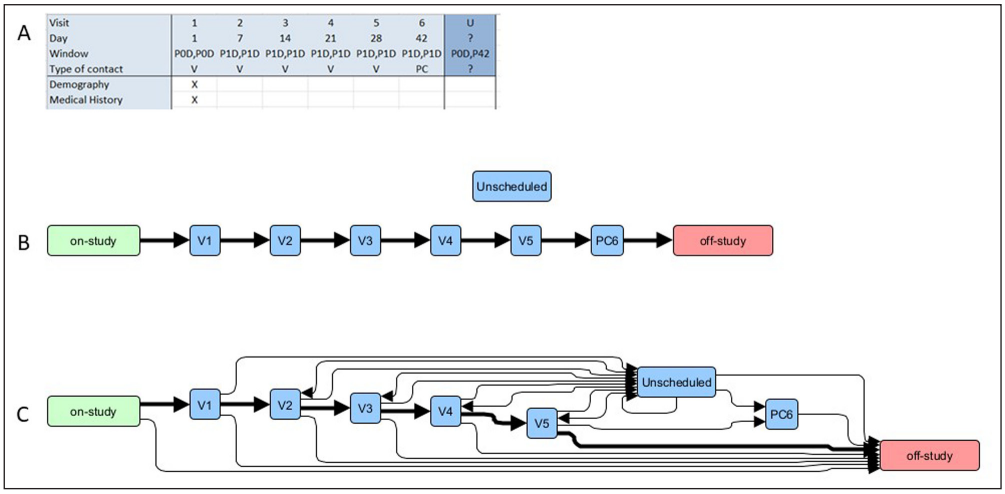

Figure 3: Expansion of an example SoA that shows **(A)**, the tabular representation from the protocol (the visits show only the first two activities); and **(B)**, a conversion to a graph representation with two additional nodes that represent (1) allocation to the study ('on-study') and (2) return to standard clinical care ('off-study'). The per-protocol sequence of visits is highlighted in **bold. (C)** shows the same 'visits' but with all possible routes that a specific study participant may followed. For example, *'on-study' to 'off-study'* could result from '*informed consent not provided*'. Similarly, at each visit the possibility exists to (1) withdraw from the study (*'V' to 'off-study'*), (2) to next make an unscheduled visit (*'V' to 'Unscheduled'*), or (3) to proceed to the next scheduled visit. Unscheduled visits may (1) return to the next scheduled visit, (2) may be repeated, or (3) require withdrawal (*'Unscheduled' to 'off-study'*). See text for further details.

Table 5: SoA Model Characteristic Attributes. Minimal viable characteristic SoA model attributes found necessary (Minimal Viable (MV) Attribute = Yes) and desirable (MV Attribute = No) to reflect SoA requirements in graph form. All of the example SoAs were able to be accurately described using these attributes. The resulting SoA Model graphs were used to generate FHIR Plan- and Activity-Definitions compliant with the HL7 Vulcan Clinical Study Schedule of Activities IG.

Attribute	MV Attribute[1]	Relationship to SoA Table Protocol, etc.	Notes
SoA Model: Interaction Node Characteristic Attributes			
nodeID	Yes	n/a	Universally Unique Identifier (UUID) – implementation requirement
type	Yes = Interaction	SoA column headings	
subtype	No	Type of sponsor/subject interaction	eg, Clinic visit, Telephone call
name	Yes	Visit, Encounter, Interaction, Appointment, etc.	As provided by SoA author in SoA table or protocol text, etc.

(Contd.)

Attribute	MV Attribute[1]	Relationship to SoA Table Protocol, etc.	Notes
description	No	Description of the interaction	
plannedTiming	Yes	Visit day, Week, etc	
referenceTimepoint	Yes	Schedule t(zero)	
plannedWindow	No	Visit timing variance	If provided
plannedDuration	No	Time period the interaction activities are to be undertaken.	eg, 24 hours
fhirDefinitionalResource	No	n/a	Target FHIR Definitional Resource (PlanDefinition)
SoA Model: Activity Node Characteristic Attributes			
nodeID	Yes	n/a	Unique UUID – implementation requirement
type	Yes = Activity	SoA 'X's	
subtype	No		Can be used to categorise activities, eg, intervention, measurement, etc.
name	Yes	Activity	As provided by SoA author in SoA table or protocol text, etc.
description	No	Description of the activity	
plannedTiming	No	Of value if to schedule activities within an interaction	eg, Measure at 10:00 am on day of visit
referenceTimepoint	No		
plannedWindow	No		
plannedDuration	No		The time period the interaction is 'active' eg, 24 hours
fhirDefinitionalResource	No		Target FHIR Definitional Resource (eg, ActivityDefinition, etc.)
SoA Model: Interaction/Activity Edge Characteristic Attributes			
edgeID	Yes	n/a	UUID – implementation requirement
transitionType	No	n/a	Timing relationship between predecessor and successor nodes (Finish-to-Start, Start-to-Start, etc.)

[1] Minimal Viable Product.

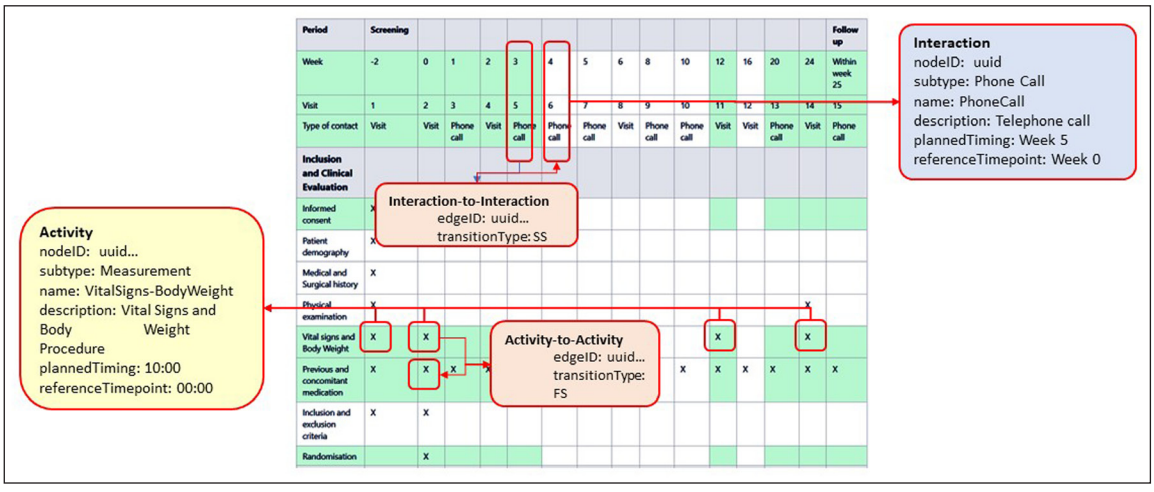

Figure 4: Graphical representation showing the relationship between a SoA table and the SoA graph model objects and characteristic attributes detailed in **Table 4**. The blue box: SoA Model Interaction; yellow box: SoA Model Activity; grey boxes: SoA Model relationships (Interaction-to-Interaction, Activity-to-Activity).

```
Instance: FHIR-SOA-Characteristics-Figure-3c
InstanceOf: StudyProtocolSoa
Description: "FHIR-SOA-Characteristics-Figure-3c"
Usage: #example

* status = #active

* title = "FHIR-SOA-Characteristics-Figure-3c"
* type = http://terminology.hl7.org/CodeSystem/plan-definition-type#clinical-protocol
* date = "2023-06-08"
* version = "V1"
* purpose = "The purpose of this PlanDefinition is to provide a worked example of Figure 3c"
* identifier[+]
  * value = "Figure-3c-Schedule-of-Activities"
  * type = http://terminology.hl7.org/CodeSystem/v2-0203#PLAC
  * use = #usual

* action[+]
  * definitionUri = "PlanDefinition/V2"
  * linkId = "V2"
  * title = "V2"
  * description = "Planned Visit [Visit-2]"
  * relatedAction[+]
    * targetId = "V1"
    * relationship = #before
    * offsetDuration = 7 'd'
    * extension[acceptableOffsetRange].valueRange.low = 1 'd'
    * extension[acceptableOffsetRange].valueRange.high = 1 'd'
  * relatedAction[+]
    * targetId = "U"
    * relationship = #before
    * offsetDuration = 1 'd'
    * extension[acceptableOffsetRange].valueRange.low = 0 'd'
    * extension[acceptableOffsetRange].valueRange.high = 7 'd'
  * relatedAction[+]
    * targetId = "OffStudy"
    * relationship = #after
    * offsetDuration = 0 'd'
    * extension[acceptableOffsetRange].valueRange.low = 0 'd'
    * extension[acceptableOffsetRange].valueRange.high = 14 'd'
```

Figure 5: FHIR Shorthand (FSH) specification of Visit-2 generated from a SoA Model graph of the SoA shown in **Figure 3c**. The SoA was initially defined as a network graph as described in the text, which was used generate appropriate FHIR resources using the python fhir.client library. The resulting JavaScript Object Notation (JSON) was confirmed as compliant with the Vulcan SoA IG. This was successfully POSTed to and retrieved from publicly available test FHIR Servers. The FSH above was created using the FSH GoFSH utility to convert the original JSON to FSH (FHIR Release Version R5). See text for further details.

In addition to the attributes described above, a *fhirDefinitionalResource* attribute was also included to provide an initial link/mapping from the SoA model object to its target FHIR Resource. These were used to define the activities that "could be performed in a time and subject-independent manner" as used by the Vulcan SoA IG. For example, the target FHIR definitional resource for a SoA interaction (visit) is a PlanDefinition.action; for an activity an ActivityDefinition.

Proof of Concept Testing

Figure 5 shows an extract from a FHIR PlanDefinition generated for the SoA in **Figure 3c** (in FHIR Shorthand [FSH] format). The example is compliant with the HL7 Vulcan Clinical Study Schedule of Activities IG[21] studyProtocolSoa profile.

The extract shows '**Visit 2'** (as a PlanDefinition.action) together with its successors and predecessors (as PlanDefinition.action.relatedActions). It shows that that **V2** is linked (relatedAction) to **V1** (the targetId), and that **V1** occurs **7 days** (offsetDuration) **before** (relationship) **V2** A visit window in the range +/− **1 day** (acceptableOffsetRange) (ie, **6–8 days**) is defined. The known possibilities – that **V2** may be followed by an unscheduled visit (**U**) or to a leave study state (**offStudy**) – are also defined. Similar FHIR PlanDefinitions can be generated from the activity 'strings' (**Figure 2**) using the SoA IG PlannedStudyVisitSoa profile (not shown).

Discussion

This work was undertaken to identify a minimum set of SoA attributes to enable study protocol requirements to be simply and accurately available in machine-readable formats, and thereafter to be accurately manipulated and generate SoAs in FHIR formats. With the FHIR resources being actively investigated to communicate study schedules and activity requirements directly to EHR systems, a minimum SoA attribute set is required for consistent basic interoperability if this approach is to be widely adopted.[8,11,22–26]

The development approach was used to ensure that any preconceived ideas about minimum requirements did not influence a systematic considered analysis. This was particularly helpful in the extraction of common SoA elements, which varied considerably in the level of information available, and how this information was presented (**Tables 2, 3** and **4**). Although the final attribute set in **Table 5** may appear somewhat obvious to the informed user, the need to ensure that each of these basic elements can be identified and are available was found to be essential to ensure robust machine-readable SoA implementations.

Graph database methods were found to offer the most robust implementation approach. **Figure 5** shows the basic relationship between protocol SoA table elements and the graph model objects. The simple correspondence between the protocol SoA and the model helped with both defining the model and confirming that it reflected accurately the protocol specifications. Coding to generate FHIR PlanDefinitions directly from the SoA model graphs is straightforward, and, following loading and recovering from publicly available FHIR servers, the graphs could be re-generated with no information loss. Representation as directed graphs also made it easy to recognise and to define other routes through a study implied in SoAs but not recognised formally (eg, the Unscheduled and Withdrawn visits [**Figures 3c** and **5**]).

Table 4 illustrates the wide range of use cases dependent upon the SoA, and hints at what additional information is required before operational implementation of each dependent activity can be finalised. This is particularly true for 'upstream' requirements, ie, those to be undertaken at investigatory sites. SoAs are intended to ***communicate study requirements*** (timings, evaluations, measurements, interventions etc.) to study teams and others in order that they may be ***operationally implemented***. The SoA model presented here offers the potential to define these requirements more systematically, and ongoing work is now centred around reviewing how it can be extended to support more specific use cases (not shown).

The minimal set of SoA characteristic attributes identified here is generic, ie, not model specific. Other SoA initiatives that focus on different primary use cases (eg the Digital Data Flow (DDF) Unified Study Definitions Model (USDM) model)[27] use alternative definitional and implementation methods. The work here may therefore have additional value for comparing and contrasting other SoA models to confirm the accuracy of conversion mapping and transformations, particularly where conversion between different models is operationally desirable or required.

Conclusion

A minimum viable set of SoA characteristics, which are able to describe common study timing requirements for defining, creating, and confirming study specific FHIR resources, have been developed, and have been shown to be able to be used to accurately create SoAs in FHIR format. The exercise has also shown that current protocol SoAs do not lend themselves easily to automate study requirements. Automation would benefit from more SoA standardisation, supported, perhaps, by a new clinical data management role of "clinical data logistician" to add the necessary study specificity. The characteristic attributes developed here should also be present and identifiable in other SoA models, and the findings may therefore have a broader applicability for confirming machine-readable SoA implementations.

Acknowledgements

The examples used here are from the Lilly Xanomeline Clinical Study Protocol (LZZT) which is a publicly available example study resource available to support operational proof-of-concept and similar investigations (also known as the CDISC Pilot Study).

The author thanks Patrick Genyn for his comments and suggestions during the preparation of this manuscript.

Competing Interests

Andrew Richardson is a Director of Zenetar Ltd., a consultancy company that provides clinical operations and related services, and is an active member of the HL7 Vulcan Schedule of Activities project and other HL7 initiatives.

Author Approval

The author has read and approved this work.

References

1. **TransCelerate.** Clinical Content & Reuse Solutions. Accessed August 3, 2023. https://www.transceleratebiopharmainc.com/assets/clinical-content-reuse-solutions/
2. *Common CSR Template: A Consistent Approach to Writing Compliant CSRs With Ease*. Transcelerate Biopharma, Clinical Content & Reuse Solutions, www.transceleratebiopharmainc.com, 2004.
3. **Parab AA, Mehta P, Vattikola A,** et al. Accelerating the adoption of eSource in clinical research: a Transcelerate point of view. *Ther Innov Regul Sci*. 2020; 54(5): 1141–1151. DOI: https://doi.org/10.1007/s43441-020-00138-y
4. **Coorevits P, Sundgren M, Klein GO,** et al. Electronic health records: new opportunities for clinical research. *J Intern Med*. 2013; 274(6): 547–560. DOI: https://doi.org/10.1111/joim.12119
5. **US Food and Drug Administration.** Real-world data: assessing electronic health records and medical claims data to support regulatory decision-making for drug and biological products. Published September, 2021. Accessed August 3, 2023. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory
6. **European Medicines Agency.** Use of real-world evidence in regulatory decision making – EMA publishes review of its studies. Published June 23, 2023. Accessed August 3, 2023. https://www.ema.europa.eu/en/news/

use-real-world-evidence-regulatory-decision-making-ema-publishes-review-its-studies
7. **Facile R, Muhlbradt EE, Gong M,** et al. Use of clinical data interchange standards consortium (CDISC) standards for real-world data: expert perspectives from a qualitative delphi survey. *JMIR Med Inform*. 2022; 10(1): e30363. DOI: https://doi.org/10.2196/30363
8. **Richardson A.** Approaches to schedule of activities (SOA) specification for automated implementation. Presentation ML05, PHUSE EU Connect 2020. https://phuse.s3.eu-central-1.amazonaws.com/Archive/2020/Connect/EU/Virtual/PRE_ML05.pdf
9. **Chatterjee A, Pahari N, Prinz A.** HL7 FHIR with SNOMED-CT to achieve semantic and structural interoperability in personal health data: a proof-of-concept study. *Sensors*. 2022; 22(10). DOI: https://doi.org/10.3390/s22103756
10. **Dridi A, Sassi S, Chbeir R, Faiz S.** A flexible semantic integration framework for fully-integrated EHR based on FHIR standard. *ICAART 2020 – Proceedings of the 12th International Conference on Agents and Artificial Intelligence*. 2020; 2: 684–691. DOI: https://doi.org/10.5220/0008981506840691
11. **Richardson A, Hummerston D.** Generating schedule of activities as FHIR resources. Paper TH07, PHUSE EU Connect 2022.
12. **HL7 Fast Healthcare Interoperability Resources.** Index – FHIR v6.0.0-Continuous Integration build. Accessed August 3, 2023. https://build.fhir.org/
13. **Wikipedia.** Use case. Accessed August 3, 2023. https://en.wikipedia.org/wiki/Use_case
14. **Product Plan.** User story examples in product development. Accessed August 3, 2023. https://www.productplan.com/glossary/user-story/
15. **Python.org.** Accessed August 3, 2023. https://www.python.org/
16. **NetworkX.** NetworkX documentation. Accessed July 26, 2023. https://networkx.org/
17. pandas – Python data analysis library. Accessed July 26, 2023. https://pandas.pydata.org/
18. **Python Package Index.** fhir.resources. Accessed July 26, 2023. https://pypi.org/project/fhir.resources/
19. **HL7 Fast Healthcare Interoperability Resources.** FHIR Shorthand – Version 3.0.0 Accessed August 3, 2023. https://build.fhir.org/ig/HL7/fhir-shorthand/
20. **yWorks.** yEd – Graph Editor. Accessed July 26, 2023. https://www.yworks.com/products/yed
21. **HL7 Fast Healthcare Interoperability Resources.** Clinical study schedule of activities. HL7.FHIR.UV.VULCAN-SCHEDULE\Home – FHIR v4.0.1. Accessed July 26, 2023. https://build.fhir.org/ig/HL7/Vulcan-schedule-ig/
22. **Low G, Ward M, Richardson A.** Study designs using FHIR: schedule of activities exchange using FHIR resources. Presentation RW05, PHUSE EU Connect 2021 https://phuse.s3.eu-central-1.amazonaws.com/Archive/2021/Connect/EU/Virtual/PRE_RW05.pdf
23. **Matsumura Y, Hattori A, Manabe S,** et al. A strategy for reusing the data of electronic medical record systems for clinical research. Stud *Health Technol Inform*. 2016; 228: 297–301.
24. **Garza M, Myneni S, Fenton SH,** et al. eSource for standardized health information exchange in clinical research: a systematic review of progress in the last year. *Journal of the Society for Clinical Data Management*. 2021; 1(2). DOI: https://doi.org/10.47912/jscdm.66
25. **Kim D, Labkoff S, Holliday SH.** Opportunities for electronic health record data to support business functions in the pharmaceutical industry – a case study from Pfizer, Inc. *Journal of the American Medical Informatics Association*. 2008; 15(5): 581–584. DOI: https://doi.org/10.1197/jamia.M2605
26. **Zong N, Wen A, Stone DJ,** et al. Developing an FHIR-based computational pipeline for automatic population of case report forms for colorectal cancer clinical trials using electronic health records. *JCO Clin Cancer Inform*. 2020; 4(4): 201–209. DOI: https://doi.org/10.1200/CCI.19.00116
27. **Clinical Data Interchange Standards Consortium.** Digital data flow. Accessed August 8, 2023. https://www.cdisc.org/ddf

How to cite this article: Richardson A. A Case Study to Identify Required Schedule of Activity Characteristic Attributes. *Journal of the Society for Clinical Data Management*. 2024; 4(1): 7, pp. 1–12. DOI: https://doi.org/10.47912/jscdm.266

Submitted: 09 August 2023 **Accepted:** 18 April 2024 **Published:** 17 July 2024

Journal of the Society for Clinical Data Management is a peer-reviewed open access journal published by Society for Clinical Data Management.

OPEN ACCESS